The
PHILOSOPHICAL
MAGAZINE

FIRST PUBLISHED IN 1798

UNIVERSITY OF HAWAII
LIBRARY
MAR 16 '55

. 46 SEVENTH SERIES No. 373 February 1955

A Journal of Theoretical Experimental and Applied Physics

EDITOR
PROFESSOR N. F. MOTT, M.A., D.Sc., F.R.S.

EDITORIAL BOARD
SIR LAWRENCE BRAGG, O.B.E., M.C., M.A., D.Sc., F.R.S.
SIR GEORGE THOMSON, M.A., D.Sc., F.R.S.
PROFESSOR A. M. TYNDALL, C.B.E., D.Sc., F.R.S.

PRICE 15s. 0d.
Annual Subscription £8 0s. 0d. payable in advance

AND PUBLISHED BY TAYLOR & FRANCIS LTD., RED LION COURT, FLEET ST., LONDON, E.C.4.

ATOMIC ENERGY

A SURVEY

by

PROFESSOR J. ROTBLAT

Published April 1954

Given as a course of lectures at the University of London under the auspices of the Atomic Scientists' Association *and the Department of Extra-Mural Studies*

CONTENTS

Atomic Research at Harwell

Speaker: Sir JOHN D. COCKCROFT, K.C.B., SC.D., LL.D., F.R.S.

Chairman: Sir RODERIC HILL, K.C.B., M.C., A.F.C., M.A.

Atomic Weapons

Speaker: Professor O. R. FRISCH, O.B.E., D.SC., F.R.S.

Chairman: Professor P. M. S. BLACKETT, M.A., D.SC., F.R.S.

Power from Atomic Energy

Speaker: Professor F. E. SIMON, C.B.E., PH.D., F.R.S.

Chairman: Sir CHRISTOPHER HINTON, M.I.C.E., M.I.MECH.E.

Radiation Hazards of Atomic Energy

Speaker: Dr. J. F. LOUTIT, M.D., M.R.C.P.

Chairman: Sir ERNEST ROCK CARLING, F.R.C.S., F.R.C.P.

Medical Uses of Atomic Energy

Speaker: Dr. E. E. POCHIN, M.D., F.R.C.P.

Chairman: Sir HENRY DALE, O.M., G.B.E., F.R.C.P., F.R.S.

Atomic Energy and Moral Issues

Speakers: Professor KATHLEEN LONSDALE, D.SC., F.R.S.
Sir GEORGE P. THOMSON, D.SC., LL.D., F.R.S.

Chairman: Professor A. V. HILL, C.H., O.B.E., SC.D., F.R.S.

Paper bound: 4/9 (70 cents) Cloth bound: 6/9 ($1)
(including postage)

Printed and Published by
TAYLOR & FRANCIS, LTD.,
RED LION COURT, FLEET STREET, LONDON, E.C.4

PERGAMON PRESS

242 MARYLEBONE ROAD · N.W.1

New Books:

Proceedings of the 1954 Glasgow Conference on Nuclear and Meson Physics

Edited by E. H. BELLAMY and R. G. MOORHOUSE

An up-to-date survey of the field, with 102 contributions on: nuclear forces and nucleon scattering; nuclear data and nuclear models; photodisintegration; beta and gamma transitions; π-mesons; field theory; high energy experimental technique; and heavy mesons and hyperons.

Royal 8vo. About 350 pages. 63s.

Progress in Nuclear Physics, Vol. 4

Edited by O. R. FRISCH

With contributions by E. W. TITTERTON, E. BREITENBERGER, R. H. DALITZ, T. G. PICKAVANCE, W. E. BURCHAM, D. WALKER, J. O. NEWTON, H. A. C. MCKAY and J. MILSTED, and D. J. HUGHES.

Royal 8vo. About 400 pages. 119 illustrations. 70s.

Isotope Geology*

By K. RANKAMA

A comprehensive introduction to the chemistry and physics of the isotopes and their natural science, treating both theory and applications, and of interest not only to the geologist but also to the physicist, chemist and biologist.

Large Royal 8vo. 552 pages. 41 illustrations. 75s.

The Theory of Cohesion†

By M. A. JASWON

" It is an excellent book for all who seek a proper understanding of the solid state."
A. H. Cottrell.

Demy 8vo. 253 pages. 42 illustrations. 37s. 6d.

Linear Feedback Analysis*

By J. G. THOMASON

A book to bridge the gap between the treatises on the practical side of electronics and those on circuit analysis, with a wealth of practical detail.

Demy 8vo. About 370 pages. 235 illustrations. About 55s.

**Distributed in the Western Hemisphere by the McGraw-Hill Book Company.*

†Distributed in the U.S. by Interscience Publishers.

XVI. *Scattering of Electrons by Charged Dislocations in Semiconductors*

By W. T. Read, Jr.
Bell Telephone Laboratories, Inc., Murray Hill, N. J.*

[Received July 1, 1954]

Abstract

Experiments suggest that dislocations in germanium act as rows of closely spaced acceptor centres. In n-type material a dislocation accepts electrons and becomes a line of negative charge surrounded by a cylindrical region of fixed positive space charge. An array of parallel dislocations will have relatively little effect on electron mobility parallel to the dislocations. The mobility normal to the dislocations will be reduced because: (1) The space charge cylinders deflect electrons and scatter the electron momentum normal to the dislocations; this reduces mean free time. Dislocation scattering is characterized by a mean free path that depends on direction. (2) Even when the dislocations are far enough apart that the mean free time is not affected, the macroscopic mobility may be reduced because the electrons cannot drift parallel to the applied field; instead they follow curved paths winding between the space charge cylinders (which in typical cases are the order of a micron in diameter).

This paper deals with both drift and Hall mobility when the current flows normal to the dislocations and a magnetic field acts (1) parallel to the dislocations and (2) normal to the dislocations.

§ 1. Introduction

The experiments of Pearson, Read and Morin (1954) indicate that dislocations in germanium act like rows of closely spaced acceptor centres. In n-type material these acceptors accept electrons from the conduction band and the dislocation becomes a line of negative charge. Surrounding the dislocation is a cylinder of fixed positive space charge. The purpose of this paper is to study the effect of these charged dislocations on the motion of conduction electrons.

In a previous paper, the author (1954) discussed the geometry of dislocations in the diamond cubic structure and described the acceptor type behaviour of dislocations by a simple model (due to Shockley 1953) in which the dislocation acceptors are identified with the dangling bonds on the edge of the extra plane of a dislocation having an edge component (that is, a dislocation which is not parallel to its slip, or Burgers vector). The electrostatic energy in the space charge region was calculated as a function of the fraction f of the dislocation acceptors which accept electrons and acquire a charge q, where q is the electronic charge. Approximate formulas were derived giving f as a function of temperature for given values of the chemical donor and acceptor concentrations and the spacing c between the acceptor centres on the dislocation. Several experiments were suggested to test the theory. The simplest array of dislocations on which to make experiments is the array produced by plastic bending. If the bending is done at a high enough temperature that there are no residual macroscopic stresses, then the dislocations are all parallel (and parallel to the axis of bending) and the density of dislocations N per cm^2 is proportional to the curvature. Such known and simple arrays were introduced by Gallagher (1952) and Pearson, Read and Morin (1954) and observed by the etch pit technique of Vogel, Pfann, Corey and Thomas (1952).

---

* Communicated by the Author.

As discussed by the author (1954), the easiest way to test the predicted f vs. T relation is to measure the conductivity parallel to the dislocations, so that scattering by the dislocations is a minimum. In this paper, however, we are interested in scattering by the dislocations and we consider the case where the current is flowing normal to the dislocations. We also deal with the Hall effect when the magnetic field is (1) parallel to the dislocations and (2) normal to the dislocations.

The dislocations affect conductivity and Hall angle because :

(1) The average concentration of conduction electrons is reduced by the number of electrons accepted by dislocations. The electrons in dislocation acceptors cannot carry current in a direction normal to the dislocations (and perhaps not in any direction).

(2) Conduction electrons are scattered by the space charge region surrounding the dislocation. Thus the mean free time is reduced.

(3) The current streamlines are distorted by the space charge cylinders ; the carriers cannot drift in straight lines parallel to the applied field ; they have to follow curve paths that wind between the cylinders. Even when the distance between dislocations is large compared to a mean free path (so that the dislocations have a negligible effect on mean free time), the dislocations may have a pronounced effect on apparent, or measured, mobility because of the distorted streamlines.

Sections 3 through 5 deal with the case where (2) above can be neglected. This will usually be the case in germanium at room temperature—see Pearson, Read and Morin (1954). Here the problem is one of flow around obstacles and a continuum rather than a particle approach is appropriate. Sections 6 through 12 deal with the effect of dislocations on the mean free times for both drift and Hall mobility ; the theory is worked out by extending Chapter 11 of Shockley (1950).

§ 2. DEFLECTION OF ELECTRONS BY THE SPACE CHARGE CYLINDER

The electrostatic potential ψ around a dislocation was discussed in the previous paper (Read 1954). The electrostatic energy $-q\psi$ of an electron increases in going toward the dislocation. Consider an electron that is moving in a straight line toward the dislocation. When it enters the space charge region it experiences a field opposing its motion. It will penetrate the space charge cylinder to a distance where $-q\psi$ is equal to the initial kinetic energy $\sim kT$ of the electron. As in the previous paper and in the theory of the abrupt *pn* junction, we shall assume that the space charge region has a sharp boundary. This is a good approximation when the potential ψ rises by kT/q in a distance that is small compared to the radius R of the space charge cylinder, as will be the case at low enough temperature. Thus an electron that strikes the cylinder head-on will penetrate only a negligible distance (compared to R). Electrons of glancing incidence will penetrate even shorter distances. We can therefore regard the electrons as making elastic collisions with rigid cylinders.

As discussed in the previous paper (Read 1954) the radius R of the cylinders is related to the spacing c between neighbouring acceptors on the dislocation, the spacing a between accepted electrons, and the excess concentration of chemical donors N_d-N_a by the condition of electrical neutrality. The number of electrons per unit length along the dislocation is $1/a=f/c$. This must be equal to the number $\pi R^2(N_d-N_a)$ of excess donors per unit length of the space charge cylinder. Thus

$$f=\pi R^2 c(N_d-N_a). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.1)$$

In the previous paper we discussed a particular type of dislocation in germanium (shown in fig. 1(*b*) of that paper), for which $c=4$ Å. In 1·7 ohm cm *n*-type germanium ($N_d-N_a=10^{15}$ cm^{-3}), f was found to be of the order of 1/10 in the temperature range $0<T<150°$. This was based on a dislocation acceptor energy level 0·225 ev below the conduction band—a value which is consistent with the experiments. For these values $R=2830$ Å, which is large compared to the spacing $a=c/f=40$ Å between electrons on the dislocation. It is, however, only about 3 times the spacing $(N_d-N_a)^{-3}$ between excess chemical donors. Thus the boundary of the space charge cylinder will not be smooth. The roughness of the boundary will give some scattering of electron momentum parallel to the dislocations. As discussed in the previous paper such scattering can be measured by measuring conductivity and Hall angle with the current parallel to the dislocations. The scattering (by dislocations) of momentum normal to the dislocations will be much greater. In what follows we shall assume that it is enough greater than dislocation scattering of momentum parallel to the dislocations can be neglected.

Let the dislocations run in the y direction, and let N be the number per cm^2 cutting the xz plane. If the specimen is bent into a radius of 5 cm. as in the experiments of Pearson, Read and Morin (1954), the density N of dislocations is of the order of 5×10^6 cm^{-2}. The exact value will depend on the orientation of the slip systems—see, for example Nye (1953) or Read (1953, p. 38–39).

§ 3. EFFECT OF DISLOCATIONS ON CURRENT STREAMLINES

In this section we consider the case where the mean free path is small compared to R. The reduction in conductivity then comes from the effects (1) and (3) discussed in § 1. We consider these in order.

Carrier Concentration. The electrons in dislocation acceptors cannot contribute to conductivity normal to the dislocations. Let n be the electron concentration in the normal (electrically neutral) n-type material and let $\epsilon=\pi R^2 N$ be the fraction of the volume occupied by the space charge cylinders. Then the average electron concentration $\langle n \rangle$ is given by

$$\langle n \rangle = n(1-\epsilon). \qquad \qquad (3.1)$$

From the definition $\epsilon=\pi R^2 N$ and the neutrality condition

$$\pi R^2(N_d-N_a)=f/c$$

we have

$$\epsilon=\frac{Nf}{c(N_d-N_a)}=\left(\frac{N_s}{N_d-N_a}\right)f \qquad \qquad (3.2)$$

where $N_s=N/c$ is the number of dislocation acceptors per cm^3. Thus, in a sample of known impurity concentration and with a known array of dislocations (which can be introduced by controlled bending and observed by the etch pit technique of Vogel, Pfann, Corey and Thomas), the fraction f of acceptors that are full can be found from ϵ. We now consider the measurement of ϵ.

Distortion of Streamlines. The $\langle n \rangle$ conduction electrons per cm^3 cannot drift in the direction of the applied field ; they have to follow curved paths that wind around the space charge cylinders. Such distortion of the streamlines reduces the apparent mobility ; that is, it reduces the measured conductivity by more than the ratio $\langle n \rangle/n$.

Let the applied field be in the x direction. Then the average field and current will also be in the x direction. The local field and current will vary throughout the specimen ; the current vanishing in the space charge cylinders where $n=0$. The local field $\mathbf{E}$ and current density $\mathbf{I}$ are related by $\mathbf{I}=q\mu n\mathbf{E}$ where μ is the drift mobility of electrons in the n-type material. (We assume that hole concentration is negligible compared to n.)

Let $\langle\,\rangle$ refer to an average over the entire specimen. Then $\langle I_x \rangle = q\mu\langle nE_x \rangle$, where $n=0$ in the space charge cylinders (which occupy a fraction ϵ of the total volume). Thus

$$\langle I_x \rangle = q\mu n(1-\epsilon)\langle E_x \rangle_n = q\mu\langle n \rangle\langle E_x \rangle_n$$

where $\langle\,\rangle_n$ refers to an average just over the normal n-type material. The measured current is $\langle I_x \rangle$ and the applied field is $\langle E_x \rangle$. Thus the measured conductivity $\sigma^*=\langle I_x\rangle/\langle E_x\rangle$ is related to the conductivity $\sigma=q\mu n$ of the n-type material by

$$\begin{aligned}\sigma^* &= \sigma\frac{\langle n\rangle}{n}\,g(\epsilon)\\ &= \sigma(1-\epsilon)g(\epsilon) \quad . \; . \; . \; . \; . \; . \; . \; . \quad (3.3)\end{aligned}$$

where $g(\epsilon)$ is a factor that measures the relative contribution of the field in the neutral n-type material to the average field. For example, if the space charge cylinders were touching (so as to form a close packed, hexagonal array) then no current could flow and all of the voltage drop would occur in the space charge cylinders and none in the n-type material. By definition

$$g(\epsilon)=\frac{\langle E_x\rangle_n}{\langle E_x\rangle}. \quad . \; . \; . \; . \; . \; . \; . \quad (3.4)$$

As $\epsilon\to 0$, $g(\epsilon)$ approaches unity. When the space charge cylinders get closer together, so that ϵ increases, then $g(\epsilon)$ decreases. Finally when the space charge cylinders are touching $g(\epsilon)=0$.

S. P. Morgan (private communication) has shown that for small ϵ, $g(\epsilon)=1-\epsilon$. Thus the apparent mobility $\sigma^*/q\langle n\rangle$ and the average electron concentration $\langle n\rangle$ decreases with ϵ at the same rate. For larger values of ϵ, Morgan has found $g(\epsilon)$ by cutting circular holes in strips of resistance paper and measuring the conductance. He has shown that, if the holes form a hexagonal array, the conductivity is isotropic. Figure 1 is a plot of his experimental results for the hexagonal array.†

If we can measure (1) the conductivity σ^* of a sample into which dislocations have been introduced by controlled plastic deformation, and (2) the conductivity σ of the sample before deformation (or the conductivity of a control), then from the ratio σ^*/σ, eqn. (3.3), and the curve of fig. 1 we have ϵ. From ϵ and eqn. (3.2) we find f. Thus the f vs. T curve can be plotted and compared with the theoretical curves to test the theory and determine accurately the energy level of the dislocation acceptors, as discussed in the previous paper.

§ 4. MAGNETIC FIELD PARALLEL TO THE DISLOCATIONS

In this section we consider the same case as in the preceding section except that, in addition to the electric field E_x, we have a magnetic field H_y parallel to the dislocations and a Hall field E_z normal to the dislocations and of such magnitude that no current flows in the z direction. The current density $\mathbf{I}$ at any point is then related to $\mathbf{E}$ and $\mathbf{H}$ by

$$\mathbf{I}=q\mu n\mathbf{E}-\frac{\mu_H}{c}\mathbf{I}\times\mathbf{H} \quad . \; . \; . \; . \; . \; . \; . \; . \quad (4.1)$$

---

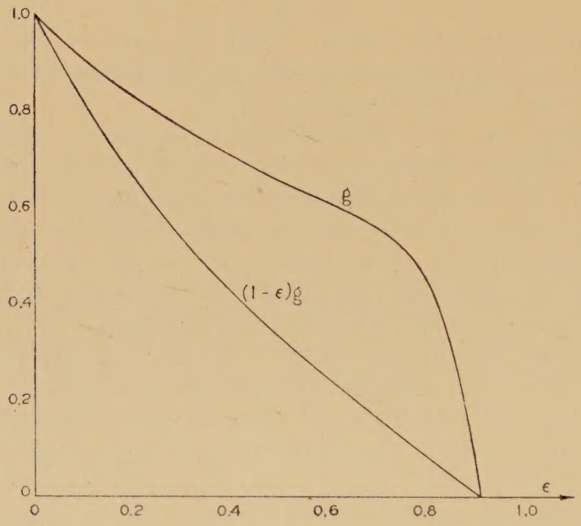
† The polar plot of conductivity in the xz plane is an ellipse. Therefore the conductivity is isotropic in the xz plane if the parallel dislocations form a regular array having two or more fold symmetry. The $g(\epsilon)$ plot however would be different for different symmetrical arrays.

where μ_H is the Hall mobility and c is the velocity of light. (Not to be confused with the c which is the spacing between acceptors along the dislocation ; the two c's never occur in the same formula and the velocity of light occurs only in the quotient H/c.) When the mean free time is determined by thermal scattering, then the drift mobility μ and the Hall mobility μ_H are simply related by

$$\mu = \frac{8}{3\pi}\mu_H. \qquad \ldots \ldots \quad (4.2)$$

See, for example, Shockley (1950), p. 278. Shockley (p. 287) also treats the more general case where both thermal and impurity scattering affect the mean free time.

Fig. 1



Variation of the dimensionless factor g and $(1-\epsilon)g$ with the fraction ϵ of space occupied by the space charge cylinders when the cylinders form a hexagonal array.

We now average eqn. (4.1) over the entire volume of the specimen. From the x components $\langle I_x\rangle = q\mu\langle nE_x\rangle$ as in the preceding section and we again have eqn. (3.3). The z components in eqn. (4.1) give $\mu_H H_y/c\langle I_x\rangle = q\mu\langle nE_z\rangle$. The x and z directions are both normal to the dislocations. Unless there is some marked anisotropy in the arrangement of dislocations we can take $\langle nE_z\rangle/\langle E_z\rangle = \langle nE_x\rangle/\langle E_x\rangle = \langle n\rangle g(\epsilon)$. This will be a good approximation if the (parallel) dislocations are arranged either randomly or in some regular array having two or more fold symmetry (so that conductivity is isotropic in the xz plane). We thus have $\mu_H H_y/c\langle I_x\rangle = q\mu\langle n\rangle\langle E_z\rangle g(\epsilon)$. When we divide this by

$$\langle I_x\rangle = q\mu\langle n\rangle\langle E_x\rangle g(\epsilon)$$

to obtain the measured Hall angle $\theta = \langle E_z \rangle / \langle E_x \rangle$, the factor $g(\epsilon)$ drops out and we have the usual formula

$$\theta = \frac{\mu_H H_y}{c}. \qquad (4.3)$$

Thus we reach the conclusion that an array of parallel dislocations should have no effect on the measured Hall angle when the electric fields are normal to the dislocations and the magnetic field is parallel to the dislocations. This provides a sensitive test to determine whether the dislocations in a bent sample are really parallel.

§ 5. MAGNETIC FIELD NORMAL TO THE DISLOCATIONS

Consider the situation discussed in § 4 but let the magnetic field be in the z direction (normal to the dislocations) and the Hall field, in the y direction (parallel to the dislocations). In this section we shall assume that the electrons in the dislocation acceptors give a negligible contribution to the conductivity in the y direction. As discussed in the previous paper, this question remains to be investigated. We would expect that an electron moving from acceptor site to acceptor site along the dislocation would have a considerably lower mobility than a conduction electron in normal n-type material. Thus, if the density fN_s of electrons in dislocation acceptors is no larger than the density of conduction electrons, we would expect the conductivity parallel to the dislocations to be due mainly to conduction electrons in the n-type material. In the bound range of temperature, where $n \ll N_d - N_a$, the dislocations might make a significant contribution to conductivity. This remains to be studied experimentally by measurements of conductivity and Hall angle vs. T at low temperature.

Consider the variation of the electric field **E** throughout the sample : **E** is constant in the y direction ; therefore the component E_y must be constant in the x and z directions as well. Then taking averages in (4.1) we have, for the y components,

$$q\mu \langle n \rangle E_y = \mu_H \langle I_x \rangle \frac{H_z}{c}$$

while the x components give $q\mu \langle n \rangle \langle E_x \rangle g(\epsilon) = \langle I_x \rangle$ as before. Thus the Hall angle θ is

$$\theta = \frac{\mu_H H_z}{c} g(\epsilon). \qquad (5.1)$$

The dislocations therefore reduce the Hall angle. Measurements of the Hall angle before and after dislocations are introduced should therefore give $g(\epsilon)$ from which f can be found as described in § 3.

The measured mobility is $\sigma^* = q\mu \langle n \rangle g(\epsilon)$ independently of the magnetic field. If we divide σ^* by θ to get $\langle n \rangle$, the factor $g(\epsilon)$ drops out and we have $\langle n \rangle$ directly in terms of quantities that are known independently of the distortion of the streamlines.

$$\langle n \rangle = \frac{\mu_H}{\mu} \frac{H_z}{qc} \frac{\sigma^*}{\theta} . \qquad (5.2)$$

This is the usual formula for average carrier concentration.

Pearson, Read and Morin (1954) used the standard formulas, eqn. (4.3) and $\mu_H = c\theta/H_z$, to calculate μ_H and $\langle n \rangle$ with the current and magnetic field both normal to the dislocations. Thus the calculated values of $\langle n \rangle$ are correct even though the mobilities are wrong by the factor $g(\epsilon)$.

§ 6. EFFECT OF DISLOCATIONS ON MEAN FREE PATH

In this section we find the mean free path associated with dislocation scattering.

First consider an electron moving with velocity $\mathbf{v}$. Let $\mathbf{v}_1$ be its velocity in the xz plane, or normal to the dislocations. As before, we treat the dislocation and its space charge as a smooth cylinder impenetrable to electrons. The conduction electrons collide with the dislocation and are deflected ; $\mathbf{v}_1$ changes on collision, while the component v_y remains the same. We now find (1) the collision rate $1/\tau_c$ for a given velocity $\mathbf{v}$ and (2) the average change in velocity $\Delta\mathbf{v}$ per collision. The product of these is the average rate of change of velocity $d\mathbf{v}/dt$ of electrons having incident velocity $\mathbf{v}$.

Collision Rate. An electron will collide with any dislocation within a distance R of its path. The probability of collision per unit time is thus $2Rv_1$ times the number N of dislocations per cm^2 cutting the xz plane. Thus the collision rate per electron is

$$\frac{1}{\tau_c} = 2RNv_1 . \qquad (6.1)$$

Average Change in Velocity. Figure 2 illustrates the reflection of an incident electron having velocity $\mathbf{v}_1$ in the xz plane. The electron comes in at a height h above a plane through the dislocation and parallel to $\mathbf{v}_1$. The change $\Delta\mathbf{v}$ in velocity on collision will be equal to the change $\Delta\mathbf{v}_1$ in $\mathbf{v}_1$. Suppose that, for any given velocity $\mathbf{v}$, all values of h are equally probable. (We discuss this point in more detail later.) Then the probability (per unit θ) of an electron being deflected through an angle θ_1 is $(d/d\theta)h/R = \frac{1}{2}\cos\phi = \frac{1}{2}\sin\frac{1}{2}\theta_1$. The average change in velocity $\langle \Delta\mathbf{v} \rangle$ is $-\mathbf{v}_1\langle 1-\cos\theta_1 \rangle$, or

$$\langle \Delta\mathbf{v} \rangle = \frac{1}{R}\int_0^R \Delta\mathbf{v}(h)\,dh = -\mathbf{v}_1 \int_0^{\frac{1}{2}\pi} (1-\cos\theta_1)\frac{1}{2}\sin\tfrac{1}{2}\theta_1\,d\theta_1 = -\frac{4}{3}\mathbf{v}_1 . \qquad (6.2)$$

The time rate of change of velocity for electrons having incident velocity $\mathbf{v}$ is

$$\frac{d\mathbf{v}}{dt} = \frac{\langle \Delta\mathbf{v} \rangle}{\tau_c} .$$

Since the y component does not change we have, using eqns. (6.1) and (6.2)

$$\frac{d\mathbf{v}_1}{dt} = -\frac{\mathbf{v}_1}{\tau_1}$$

where τ_1 is the mean free time for dislocation scattering and is given by

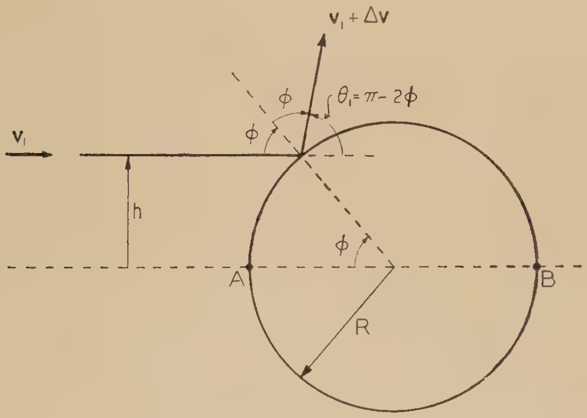
$$\frac{1}{\tau_1} = \frac{\langle 1 - \cos\theta_1 \rangle}{\tau_c} = \frac{v_1}{l_1} \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6.3)$$

where

$$l_1 = \frac{2}{3RN} \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6.4)$$

is the constant mean free path for scattering of momentum in the plane normal to the dislocations.

Fig. 2



Reflection of an electron from an impenetrable space charge cylinder.

§ 7. Basic Assumptions

In this section and the following one, we consider the combined effects of applied fields and scattering. The analysis will be an extension of Chapter 11 of Shockley (1950); the same terms are used and some of the equations are taken over directly. However, it is necessary to take a somewhat more general view from the beginning because, when dislocation scattering is included, the mean free time is not constant over an energy surface in momentum space.

This section discusses the basic assumptions peculiar to dislocation scattering. Section 8 will treat the motion of electrons in momentum space. Section 9 returns to scattering and uses some results of § 8 to treat the combined effects of dislocation and thermal scattering. The following sections consider combined thermal and dislocation scattering under various orientations of the electric and magnetic fields. The steady state nonequilibrium distribution of electrons in momentum space is found by equating the rates of change due to the fields and scattering. The approach is similar to that given by Shockley (1950) except that we consider each element of momentum space separately instead of finding the rate of change of current in an energy shell.

We base the analysis on the following two approximations :

(1) The fields **E** and **H** are constant throughout the material.

(2) The distribution in velocity of the electrons incident on a reflecting cylinder is the same as the average distribution over the whole volume of material.

Neither of these assumptions is exact but they become increasingly good approximations as the distance between dislocations increases. Assumption (1) is most nearly fulfilled when the mean free path for thermal scattering is longer than that for dislocation scattering. As thermal scattering begins to dominate, the errors in the calculation of dislocation scattering are greater ; however, they have a smaller effect on the final result. The two assumptions are not necessarily exact even when there is only one dislocation ; as the following argument will show.

Consider the distribution of electrons moving in the direction opposite to the applied electric field and net current. Let this be the $+x$ direction. The electrons are incident on the $-x$ side of the cylinder. Take the electron distribution at the point where the incidence is normal, for example point A in fig. 2. If the distribution of electrons in velocity space is the same at A as in the average for the whole specimen, then there will be an excess of electrons going in the $+x$ direction. However, the number of incident and reflected electrons are equal at A ; hence there is also an excess of electrons going in the $-x$ direction. The reverse is true on the $+x$ side of the cylinder at the point B ; B will therefore have a net electron deficiency. The same general argument applies to electrons moving in other directions and leads to the conclusion that the electron density is greater on the $-x$ side of the cylinder than on the $+x$ side (the applied field being in the x direction). This nonuniformity of electron density constitutes a line dipole, which will induce a compensating dipole by displacing the axis of the space charge cylinder in relation to the charged dislocation. The effect of the two opposite line dipoles is to produce some nonuniformity in the electric field near the cylinder.

§ 8. Motion of Electrons in Momentum Space

Sections 6 and 7 dealt with the motion of electrons in the semiconducting material. Now we deal with the motion of electrons in momentum space and the rate of change of electron distribution under electric and magnetic fields. The symbols used and some of the results are from Chapter 11 of Shockley (1950). However, the analysis is somewhat more general and is based on an approach suggested by C. Herring. As in Shockley's treatment we assume the energy surfaces are spherical.

Consider a point $\mathbf{P}$ in momentum space. Let $f(\mathbf{P})$ be the fraction of the energy states at $\mathbf{P}$ that are full at thermal equilibrium. Let $f'(\mathbf{P})$ be the fraction that are full in the nonequilibrium conditions, that is, when a current is flowing. When there is an electric field $\mathbf{E}$, then $f'=f+f_1$ where f_1 is proportional to E. Shockley treats the current (in an energy shell) associated with f_1, then adds the magnetic field $\mathbf{H}$ and finds the Hall current (current normal to $\mathbf{H}$ and $\mathbf{E}$) from the current due to $\mathbf{E}$. We cannot follow Shockley's procedure, however, since it requires that the mean free time τ be constant over an energy surface, where the magnitude v of the velocity is constant. For dislocation scattering $\tau_1=\tau_1(v_1)$. Hence, the mean free time for combined dislocation and thermal scattering (which we shall derive in § 9) is a function of both v and v_1. It is therefore constant, not over a sphere, but rather over a circle around the P_y axis.

We begin with the equation of continuity in momentum space

$$\frac{\partial f'}{\partial t}=q\left(\mathbf{E}+\frac{\mathbf{v}\times\mathbf{H}}{c}\right)\cdot\nabla_P f'. \qquad (8.1)$$

It will be understood that all time derivatives in this section refer to changes due to the fields only. Later we add these rates and the corresponding rates due to scattering and set the sum equal to zero. The resulting equations give $f'=f'(\mathbf{P})$, from which we find the currents for given electric and magnetic fields.

The general approach suggested by Herring is the following: f' is a function of $\mathbf{E}$ and $\mathbf{H}$. We expand f' in a series in powers of the magnitudes E and H. The terms in H alone vanish (H simply makes electrons flow around in circles and therefore leaves the distribution unchanged). Since we are interested in measuring conductivity and Hall angle, which are linear in the fields, we neglect second and higher powers of E and H. So

$$f'=f+f_1+f_2 \qquad (8.2)$$

where f_1 is proportional to E, as in Shockley's treatment, and f_2 is proportional to EH.

The next step is to substitute (8.2) into (8.1). First however consider the meaning of the derivatives $\mathbf{E}\cdot\nabla_P$ and $\mathbf{v}\times\mathbf{H}\cdot\nabla_P$. The first is simply proportional to the derivative in the direction of the electric field; $\mathbf{v}\times\mathbf{H}\cdot\nabla_P$ is the derivative in a direction tangent to the constant energy surface, which we shall call S, and in a plane normal to $\mathbf{H}$. Thus it is the derivative along a circle on the sphere S and around the direction of $\mathbf{H}$. The equilibrium distribution $f(\mathbf{P})=f(\mathscr{E})=f(v)$ is constant over the spherical energy surface. Using this result in (8.1) and separating the terms in E from the terms in HE and neglecting terms in E^2 and H^2 we have

$$\frac{\partial f_1}{\partial t}=q\mathbf{E}\cdot\nabla_P f \qquad (8.3)$$

and

$$\frac{\partial f_2}{\partial t}=\frac{q}{c}\mathbf{v}\times\mathbf{H}\cdot\nabla_P f_1. \qquad (8.4)$$

In the following section we find f_1 from (8.3) together with the rate of change of f_1 due to scattering. When f_1 is known f_2 is found from (8.4) and the formulas for scattering. To conclude this section we simply state the following results which are discussed more fully in § 11.4 of Shockley (1950).

$$\nabla_P f=\frac{df}{d\mathscr{E}}\nabla_P\mathscr{E}=-\frac{f(1-f)}{kT}\mathbf{v}\doteq-\frac{\mathbf{v}f}{kT} \quad . \quad . \quad . \quad (8.5)$$

where the approximation holds for non-degenerate semiconductors, for which $f \ll 1$ for electrons in the conduction band.

We now have for (8.3)

$$\frac{\partial f_1}{\partial t}=-\frac{q\mathbf{E}\,.\,\mathbf{v}}{kT}f=-\left(\frac{qE_xf}{kT}\right)v_x. \quad . \quad . \quad . \quad (8.6)$$

Thus over an energy surface (where f is constant) the rate of change of f_1 is proportional to $-v_x$.

We shall show in the next section that f_1 is also proportional to v_x. This result, however, does not necessarily extend to f_2. We shall see in § 12 that, when $\mathbf{H}$ is in the z direction, so that f_2 is associated with a flow in the y direction, then f_2 (on an energy surface) is *not* simply proportional to v_y.

§ 9. THERMAL AND DISLOCATION SCATTERING

In this section we consider combined dislocation and thermal scattering with an electric field E_x in the x direction, that is, normal to the dislocations. We shall show that

$$f_1=-\frac{qE_x\tau f}{kT}v_x \quad . \quad . \quad . \quad . \quad . \quad (9.1)$$

where the mean free time τ is given by

$$\frac{1}{\tau}=\frac{1}{\tau_t}+\frac{1}{\tau_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9.2)$$

where τ_t and τ_1 are the mean free times for thermal and dislocation scattering respectively when each acts alone. That reciprocal mean free times are additive is neither obvious nor generally true. The proof for the present case follows.

Again we use the notation and basic relations of Chapter 11 of Shockley (1950). W_{ij} is the matrix element giving the probability that an electron at $\mathbf{P}_i$ in momentum space will make a transition to another point $\mathbf{P}_j$ on the same energy surface S when all the states at $\mathbf{P}_j$ are empty. The rate of scattering from $\mathbf{P}_i$ to $\mathbf{P}_j$ is then proportional to $f_i'(1-f_j')W_{ij}$. The *net* rate of scattering from $\mathbf{P}_i$ to $\mathbf{P}_j$ is

$$[f_i'(1-f_j')-f_j'(1-f_i')]W_{ij}=(f_i'-f_j')W_{ij}=(f_{1i}-f_{1j})W_{ij}$$

(since f is constant over S).* In the cases that Shockley treats W_{ij} is a function only of the angle θ between $\mathbf{P}_i$ and $\mathbf{P}_j$. For thermal scattering alone W_{ij} is a constant W_t. The mean free time τ_t for thermal scattering is then equal to the probability of a collision ; so

$$\frac{1}{\tau_t}=\int_S W_t dS=W_t S \qquad \ldots\ldots \quad (9.3)$$

where S is the area of the spherical energy surface in momentum space.

For dislocation scattering W_{ij} vanishes except when $\mathbf{P}_i$ and $\mathbf{P}_j$ lie on the same circle C around the P_y axis (since the y component of momentum is not changed in a collision). When $\mathbf{P}_i$ and $\mathbf{P}_j$ do lie on the same circle C, then W_{ij} for dislocation scattering is equal to a function—call it W_1—of the angle θ_1 between the projections of $\mathbf{P}_i$ and $\mathbf{P}_j$ in the xz plane. This is the same θ_1 used in § 6 and illustrated in fig. 2. By definition the collision time τ_c, eqn. (6.1), is related to $W_1(\theta_1)$ by

$$\frac{1}{\tau_c}=\int_C W_1(\theta_1)dC. \qquad \ldots\ldots\ldots \quad (9.4)$$

W_1 is a probability per unit length in momentum space : W_t was a probability per unit area. We found in § 6 that $W_1(\theta)$ was proportional to $\frac{1}{2}\sin\frac{1}{2}\theta_1$. Equation (9.4) determines the constant of proportionality. The result is $W_1=1/4R\tau_c \sin\frac{1}{2}\theta_1$. The mean free time τ_1 for dislocation scattering is defined in terms of W_1 by

$$\frac{1}{\tau_1}=\frac{1}{\tau_c}\int_C (1-\cos\theta_1)W_1(\theta_1)dC \qquad \ldots\ldots \quad (9.5)$$

which gives $\tau_1=\frac{3}{4}\tau_c$.

From these definitions we now go on to prove eqns. (9.1) and (9.2).

The W's for various scattering mechanisms are additive. In the present special case this will lead to the result that the reciprocal mean free times are additive.

Consider the rate of change of $f'(\mathbf{P})$. In § 8 we used $\partial/\partial t$ for the time rate of change due to the electric and magnetic fields. We shall continue this notation. To distinguish acceleration and scattering terms we use a dot to indicate the time rate of change due to scattering. Then $f'=f_1$ is

$$\dot{f}_1(\mathbf{P})=\int_S [f_1(\mathbf{P}')-f_1(\mathbf{P})]W_t dS'+\int_C [f_1(\mathbf{P}')-f_1(\mathbf{P})]W_1(\theta_1)dC'. \quad \ldots \quad (9.6)$$

The integrals are taken over S and C with $\mathbf{P}'$ as the variable of integration. Equation (8.6) gives $\partial f_1/\partial t$, which is due to the field E_x. Setting the sum of the two rates equal to zero gives an integral equation for $f_1(\mathbf{P})$ in terms of known quantities. We now show that (9.1) is a solution to the integral equation.

---

* The same result would have been obtained by neglecting the $(1-f)$ factors. In other words, we make no error in the final result if we assume that the probability of a transition from one state to another is (a) proportional to the probability of occupation of the state that the electron is *leaving* and (b) independent of the probability of occupation of the state that the electron is *entering*. (See Shockley 1950, p. 256.) We shall use this result in § 12.

The quantities τ and f in (9.1) are constant over the circle C and over every circle around the y axis. Thus f_1 is an odd function of v_x, and therefore of P_x. So the first integral in (9.6) becomes

$$-W_t Sf_1(\mathbf{P})=-\frac{f_1(\mathbf{P})}{\tau_t}.$$

The second integral involves only the x and z components of momentum and velocity. We can express everything in terms of the projection of the velocity in the xz plane. From simple trigonometry we have $v_x'=-v_x(1-\cos\theta_1)+v_z\sin\theta_1$. Thus substituting (9.1) into the second integral in (9.6) gives

$$-\frac{qE_x\tau f}{kT}\left[v_x\int_C(1-\cos\theta_1)W(\theta_1)dC+v_z\int_C\sin\theta_1 W(\theta_1)dC\right].$$

Since $W(\theta_1)=W(-\theta_1)$, the coefficient of v_z vanishes. Then using (9.1) and (9.5) we are left with $-f_1(\mathbf{P})/\tau_1$ for the second integral in (9.6); so (9.6) becomes

$$\dot{f}_1=-f_1\left(\frac{1}{\tau_t}+\frac{1}{\tau_1}\right)=-\frac{f_1}{\tau}. \qquad (9.7)$$

Adding the scattering rate (9.7) to the acceleration rate (8.6) and setting the sum equal to zero gives (9.1).

If there were impurity as well as dislocation and thermal scattering then the reciprocal mean free times would not be additive.

§ 10. DRIFT MOBILITY

As in the preceding section let there be an electric field E_x which gives rise to a current I_x. Suppose I_x is scattered by both dislocations and thermal vibrations. In this section we find the drift mobility μ in terms of the mean free path l_t for thermal scattering and the mean free path l_1 for dislocation scattering; l_1 is given by (6.4). The mean free path for thermal scattering is derived by Shockley (1950). It can also be determined experimentally as a function of temperature by measurements on samples containing a small density of dislocations and impurities.

Consider an element of momentum space having a number ds of states. Let $d\mathbf{I}$ be the element of current density due to these states. The net current arises from $f_1=f'-f$. Thus $d\mathbf{I}=q\mathbf{v}f_1ds$. From (9.1)

$$d\mathbf{I}=\frac{q^2}{kT}v_xE_x\mathbf{v}\tau f ds.$$

Now $d\mathbf{I}$ is expressed in terms of the equilibrium distribution f. The equilibrium number of electrons in the ds states is $dn=fds$. Integrating over all of momentum or velocity space gives the total current density

$$I_x=\frac{q^2E_x}{kT}\int v_x^2\tau dn. \qquad (10.1)$$

For spherical energy surfaces the other components of current vanish. The drift mobility $\mu=I_x/qnE_x$ is found from (10.1):

$$\mu=\frac{q}{kT}\frac{\int v_x{}^2\tau\,dn}{\int dn}. \qquad (10.2)$$

It is convenient to use spherical coordinates in velocity space with

$$\left.\begin{aligned} v_1&=v\sin\theta, & v_x&=v\sin\theta\cos\phi, \\ v_y&=v\cos\theta, & v_z&=v\sin\theta\sin\phi. \end{aligned}\right\} \qquad (10.3)$$

The mean free times are related to the mean free paths by $\tau_t=l_t/v$ and $\tau_1=l_1/v_1$. The paths l_t and l_1 are constant over velocity space at a given temperature. Thus

$$\tau=\frac{1}{v\left(\dfrac{1}{l_t}+\dfrac{\sin\theta}{l_1}\right)}. \qquad (10.4)$$

The number dn of electrons in an element of velocity space is

$$dn=v^2\sin\theta\exp(-mv^2/2kT)d\theta\,d\phi\,dv.$$

It is convenient to introduce the average thermal velocity

$$\langle v\rangle=\frac{\int v\,dn}{\int dn}=\sqrt{\frac{8kT}{\pi m}}. \qquad (10.5)$$

Now (10.2) becomes

$$\mu=\frac{q}{kT}\frac{\langle v\rangle}{2}\int_0^{\pi/2}\frac{\sin^3\theta\,d\theta}{\dfrac{1}{l_t}+\dfrac{\sin\theta}{l_1}}. \qquad (10.6)$$

When there is only thermal scattering $l_1=\infty$ and the integral equals $2l_t/3$. For dislocation scattering $l_t=\infty$ and the integral is $\pi/4l_1$. It is therefore convenient to define a mean free path l_μ where

$$\frac{1}{l_\mu}=\frac{3}{l_t}+\frac{8}{\pi l_1} \qquad (10.7)$$

and set

$$\mu=\frac{q}{kT}\langle v\rangle l_\mu F_1 \qquad (10.8)$$

where F_1 is a function of the ratio l_1/l_t and is given by

$$F_1\left(\frac{l_1}{l_t}\right)=\frac{1}{2}\left(3\frac{l_1}{l_t}+\frac{8}{\pi}\right)\int\frac{\sin^3\theta\,d\theta}{\dfrac{l_1}{l_t}+\sin\theta}. \qquad (10.9)$$

The left hand side of fig. 3 shows F_1 vs. l_1/l_t for $l_1\leqslant l_t$. The right hand side shows F_1 vs. l_t/l_1 for $l_t\leqslant l_1$. As insured by the definition,

$$F_1(0)=F_1(\infty)=1.$$

Since F_1 is always between 0·99 and 1 it is sufficient for practical purposes to use the approximation $F_1=1$. In other words the reciprocal of the resultant mean free path is simply a weighted average of the reciprocal mean free paths for dislocation and thermal scattering individually.

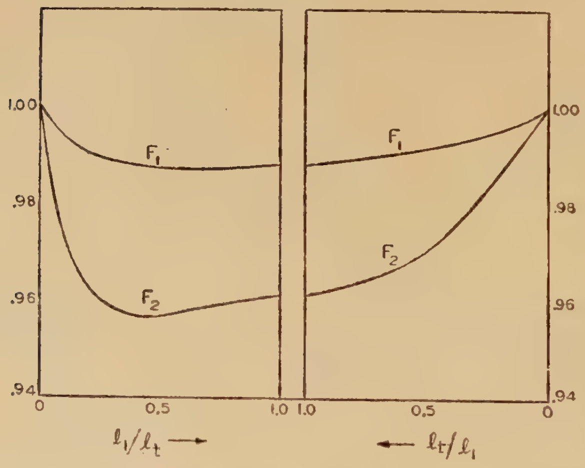
§ 11. MAGNETIC FIELD PARALLEL TO THE DISLOCATIONS

Now that we have f_1 we can proceed to find f_2. Since f_1 is uniquely determined by the field E_x, it is independent of the magnetic field **H**. In this section we take $H = H_y$ in the direction parallel to the dislocations. The procedure will be as follows: first we find $\partial f_2/\partial t$ due to the fields by substituting eqn. (9.1) for f_1 into eqn. (8.3). It will turn out that $\partial f_2/\partial t$ due to the fields has the same form as $\partial f_1/\partial t$; so the same argument (§ 9) that gave f_1 can be used to find f_2. We then evaluate the current I_z associated with f_2. From this we find the Hall angle and Hall mobility.

Substituting f_1 from (9.1) into (8.3) gives

$$\frac{\partial f_2}{\partial t} = -\frac{q^2 E_x}{kT} \mathbf{v} \times \frac{\mathbf{H}}{c} \cdot \nabla_P(\tau f v_x). \quad . \quad . \quad . \quad (11.1)$$

Fig. 3



Variation of the dimensionless factors F_1 and F_2 with the ratios l_1/l_t and l_t/l_1, where l_t is the mean free time for thermal scattering and l_1, for dislocation scattering.

As discussed in § 8 in the paragraph preceding eqn. (8.3), the operator $\mathbf{v} \times \mathbf{H} \cdot \nabla_P$ takes the derivative along the circumference of a circle on the energy surface and around the axis parallel to **H**. In this section **H** is parallel to the y axis; so the circle in question is the circle C on which the mean free time τ is constant. Then, since $f = f(v)$ is also constant on C.

$$\mathbf{v} \times \mathbf{H} \cdot \nabla_P(\tau f v_x) = \tau f \mathbf{v} \times \mathbf{H} \cdot \nabla_P v_x.$$

The vector $\nabla_P v_x$ is in the x direction and has the magnitude $1/m$ where

m is the mass of an electron (which is constant for spherical energy surfaces). Hence $\mathbf{v}\times\mathbf{H}\cdot\nabla_p v_x=-v_z H_y/m$ and (11.1) becomes

$$\frac{\partial f_2}{\partial t}=\left(\frac{\omega q E_x f \tau}{kT}\right)v_z, \qquad (11.2)$$

where $\omega=qH/mc$ is the cyclotron frequency (frequency with which electrons move around in circles in velocity space when $\mathbf{H}$ acts alone).

Equation (11.2) shows that $\partial f_2/\partial t$ due to the fields is proportional to v_z. The mean free time in the z direction is the same as that in the x direction (unless there is some marked anisotropy in the arrangement of dislocations). Thus the arguments applied to f_1 in § 9 apply also to f_2 and we have due to scattering

$$\dot{f}_2=-\frac{f_2}{\tau}. \qquad (11.3)$$

Adding the two rates in (11.3) and (11.4) and setting the sum equal to zero gives

$$f_2=\frac{\omega q}{kT}E_x f \tau^2 v_z. \qquad (11.4)$$

The current associated with f_2 is the integral of $q\mathbf{v}f_2$ over all the states in velocity space. All the components of $\mathbf{I}_2$ vanish except

$$I_z=\frac{\omega q^2 E_x}{kT}\int \tau^2 v_z^2\,dn. \qquad (11.5)$$

Again using spherical coordinates (10.3), we have

$$I_z=\frac{q^2}{kT}E_x\frac{n}{2}\int_0^{\pi/2}\frac{\sin^3\theta\,d\theta}{\left(\frac{1}{l_t}+\frac{\sin\theta}{l_1}\right)^2}. \qquad (11.6)$$

For thermal scattering alone, the integral equals $\frac{2}{3}l_t$. For dislocation scattering alone it equals l_1.

We will usually be interested in the open circuit Hall voltage. Let E_z be the applied Hall field required to give zero current in the z direction. We are assuming the x and z directions are equivalent as far as scattering is concerned ; so $I_z=q\mu n E_z$, where the drift mobility μ is given by (10.7) and I_z, by (11.6). The measured Hall angle is $\theta=E_z/E_x$. The Hall mobility μ_H is defined by

$$\mu_H=\frac{c\theta}{H_y}=\frac{cE_z}{H_y E_x}. \qquad (11.7)$$

In the present case $E_z/E_x=I_z/I_x=I_z/qn\mu E_x$. Thus μ_H is proportional to the integral in (11.6) divided by the integral in (10.6). The quotient of these integrals is l_t when there is only thermal scattering and $(4/\pi)l_1$ when there is only dislocation scattering. Thus it is convenient to define a mean free path l_H where

$$\frac{1}{l_H}=\frac{1}{l_t}+\frac{\pi}{4l_1}. \qquad (11.8)$$

Then

$$\mu_H=\frac{q}{m}\frac{l_H}{\langle v\rangle}F_2 \qquad (11.9)$$

where F_2 is a dimensionless quantity equal to the integral in (11.6) divided by l_H times the integral in (10.6). $F_2=1$ when either dislocation or thermal scattering acts alone. Figure 3 also shows the variation of F_2 with the ratios l_1/l_t and l_t/l_1. F_2 varies from a maximum of unity to a minimum of 0·956.

The Hall mean free time

$$\tau_H=\frac{\theta}{\omega}=\frac{\mu_H m}{q} \qquad . \quad . \quad . \quad . \quad . \quad . \quad (11.10)$$

is now

$$\tau_H=\frac{l_H}{\langle v\rangle}F_2. \qquad . \quad . \quad . \quad . \quad . \quad . \quad (11.11)$$

§ 12. MAGNETIC FIELD NORMAL TO THE DISLOCATIONS

In this section we take the magnetic field in the z direction, that is, normal to the dislocations. The electric field is still in the x direction. Again we have equation (11.1). Now $\mathbf{v}\times\mathbf{H}\,.\,\nabla_P$ is the derivative along a circle about the z axis. The mean free time τ is constant on a circle about the y axis. Thus the results will involve the gradient of τ in momentum space. Equation (11.1) becomes

$$\frac{\partial f_2}{\partial t}=\frac{q^2E_x}{kT}f\left[\tau\mathbf{v}\times\frac{\mathbf{H}}{c}\,.\,\nabla_P v_x+v_x\mathbf{v}\times\frac{\mathbf{H}}{c}\,.\,\nabla_P\tau\right] \quad . \quad . \quad . \quad (12.1)$$

where $\mathbf{v}\times\mathbf{H}\,.\,\nabla_P v_x=v_yH_z/m$. Now consider the second term in the brackets :

$$\mathbf{v}\times\mathbf{H}\,.\,\nabla_P\tau=-\tau^2\mathbf{v}\times\mathbf{H}\,.\,\nabla_P\left(\frac{v}{l_t}+\frac{v_1}{l_1}\right)$$

$$=-\frac{\tau^2}{l_1}\mathbf{v}\times\mathbf{H}\,.\,\nabla_P v_1=-\frac{\tau^2H_zv_yv_x}{l_1mv_1}.$$

Thus (12.1) becomes

$$\frac{\partial f_2}{\partial t}=\frac{q\omega E_x}{kT}\left(\tau-\frac{\tau^2v_x^{\,2}}{l_1v_1}\right)v_yf, \quad . \quad . \quad (12.2a)$$

where $\omega=qH_z/mc$. Clearly f_2 gives a current $I_2=I_y$ in the y direction. Since f_2 is not constant on a circle about the y axis, we cannot take $\dot{f}_2=-f_2/\tau_t$. Instead let us consider the change in I_2 due to changing occupation of the states along two equal circles about the y axis in velocity space. The circles have equal values of v, v_1 and τ but opposite values of v_y. Call the two circles together C. The fields cause a rate of change of current proportional to $\langle v_y(\partial f_2/\partial t)\rangle_c$ where the average is taken over the two circles and simply involves integrating with respect to ϕ. From (12.2a) and $\langle v_x^{\,2}\rangle_c=v_1^{\,2}/2$ we have

$$\left\langle v_y\frac{\partial f_2}{\partial t}\right\rangle_c=\frac{q\omega E_x}{kT}\left(\tau-\frac{\tau^2v_1}{2l_1}\right)v_y^{\,2}f. \quad . \quad . \quad (12.2b)$$

Scattering has the following effects on occupation of the states along the circles C :

(1) Dislocation scattering causes a redistribution of electrons on each circle.

(2) Electrons are thermally scattered *out* of the states along the circles.

(3) Electrons are thermally scattered *into* the circles.

Clearly (1) has no effect on I_2. As discussed in § 9 (footnote) we can assume that (*a*) the probability that an electron will be scattered out of a group of states is proportional to the fraction of those states that are full and (*b*) the probability that an electron will be scattered into a group of states is independent of the occupation of those states. If electrons are equally likely to be thermally scattered into any point on the circles, then (3) gives a change in current proportional to $\langle \mathbf{v} \rangle_c = 0$. The only non-vanishing contribution comes from (2) and is proportional to $-1/\tau_t \langle v_y f_2 \rangle_c$. Setting this equal and opposite to (12.2*b*) we have

$$\langle v_y f_2 \rangle_c = \frac{q\omega E_x \tau_t \tau}{kT} \left(1 - \frac{\tau v_1}{2l_1}\right) v_y^2 f. \quad . \quad . \quad . \quad (12.3)$$

Expressing this in spherical coordinates and integrating $q\langle v_y f_2 \rangle_c$ over all states in velocity space gives

$$I_y = \frac{q^2 E_x \omega n \, l_t \, l_2}{kT} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12.4)$$

where the length l_2 is defined by

$$l_2 = \int_0^{\pi/2} \frac{\cos^2\theta \sin\theta \, d\theta}{\dfrac{1}{l_t} + \dfrac{\sin\theta}{l_1}} - \frac{1}{2l_1} \int_0^{\pi/2} \frac{\sin^2 d\theta}{\left(\dfrac{1}{l_t} + \dfrac{\sin\theta}{l_1}\right)^2}. \quad . \quad . \quad (12.5)$$

When there is only thermal scattering ($l_1 = \infty$), then $l_2 = l_t/3$. As l_1/l_t approaches zero l_2 approaches zero.

There are two ways of defining the Hall mobility $\mu_H = c\theta/H_z$. First we can measure the short circuit Hall current I_y and define θ as the ratio I_y/I_x. Secondly we can apply a Hall field E_y so that no current flows in the y direction and define θ as the ratio E_y/E_x. In the present case, where the x and y directions are not equivalent, these two definitions of θ will not give the same result. We shall define $\theta = E_y/E_x$ because this is what is usually measured. The scattering in the y direction is entirely thermal. Therefore $I_y = q\mu_t n E_y$ where $\mu_t = q\langle v \rangle l_t / 3kT$ is the drift mobility for thermal scattering. Setting the I_y due to E_y equal to the I_y in eqn. (12.4) we have

$$\mu_H = \frac{c\theta}{H_z} = \frac{3q \, l_2}{m\langle v \rangle}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12.6)$$

It is convenient to express the results by the ratio of μ_H to the Hall mobility $\mu_{Ht} = \frac{3}{8}\mu_t$ for thermal scattering alone. Then, using (10.5) we have

$$\frac{\mu_H}{\mu_{Ht}} = \frac{3l_2}{l_t}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12.7)$$

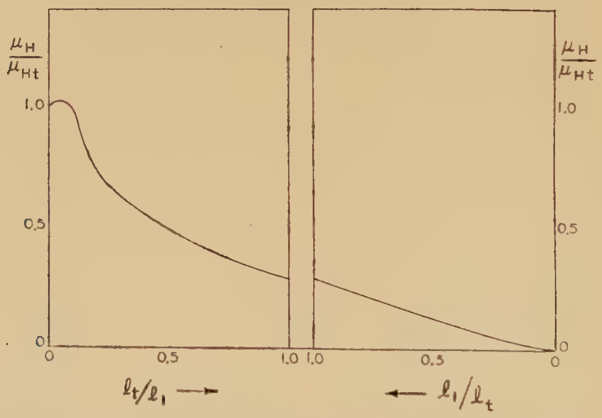
Figure 4 is a plot of μ_H/μ_{Ht} vs. l_1/l_t for $l_1 \leq l_t$ and vs. l_t/l_1 for $l_t \leq l_1$.

§ 13. CONCLUSIONS

This section reviews the range of validity of the results derived in the preceding sections and considers possible experiments and further theoretical analysis.

As discussed in § 7, the analysis of mean free time will be most accurate when the dislocations are farthest apart and the fraction ϵ of space occupied by the space charge cylinders is small. As ϵ increases, assumptions (1) and (2) in § 7 become increasingly poor approximations. Another effect that limits the validity of the results at large ϵ is the following: according to the derivation in § 9 a probability can be assigned to any given sequence of transitions, or deflections. For any such sequence we can find the positions of the deflecting cylinders. We have assumed, however, that the space charge cylinders do not overlap. Thus certain sequences of transitions are forbidden because they would require overlapping cylinders. This effect is probably not important except when the average minimum spacing between cylinders becomes less than R.

Fig. 4



Variation of the ratio of Hall mobility μ_H for combined dislocation and thermal scattering to Hall mobility μ_{Ht} for thermal scattering alone. The electric and magnetic fields are both normal to the dislocations.

To study the effect of dislocations on mean free time it is desirable not only to have ϵ small but also to have $1/l_1$ at least comparable to $1/l_t$. This requires R to be small since l_1 is proportional to R and ϵ to R^2. However, to make R small it is necessary to have a large excess donor concentration $N_d - N_a$, which means that impurity scattering becomes important. Because the effect of dislocations on mean free time is most pronounced at low temperatures, it would be useful to work out the theory of combined dislocation and impurity scattering. However, in any case, the formulas of §§ 8 to 12 with $l_t = \infty$ will be valid when dislocation scattering is the dominant scattering mechanism.

Over most of the temperature range where thermal scattering is more important than impurity scattering we have to consider both the distortion of the streamlines and the reduction in mean free time. As an approximation we may use the formulas of §§3 to 5 with the drift and Hall mobilities given by the formulas in §§8 to 12. Then, to find the effect of dislocations on mean free time, it would be best to have the magnetic field along the dislocations so that the factor $g(\epsilon)$ cancels. For the same reason to find $\langle n \rangle$ we would want the magnetic field normal to the dislocations.

ACKNOWLEDGEMENTS

The author wishes to thank C. Herring, S. P. Morgan, W. Shockley, R. C. Fletcher and G. Wannier for helpful discussion.

REFERENCES

GALLAGHER, C. J., 1952, *Phys. Rev.*, **88,** 721.
NYE, J. F., 1953, *Acta Metallurgica*, **1,** 153.
PEARSON, G. L., READ, W. T., Jr., and MORIN, F. J., 1954, *Phys. Rev.*, **93,** 666.
READ, W. T., Jr., 1953, *Dislocations in Crystals* (New York : McGraw Hill) ; 1954, *Phil. Mag.*, **45,** 775.
SHOCKLEY, W., 1950, *Electrons and Holes in Semiconductors* (New York : Van Nostrand) ; 1953, *Phys. Rev.*, **91,** 228.
VOGEL, F. L., PFANN, W. G., COREY, H. F., and THOMAS, E. E., 1953, *Phys. Rev.*, **90,** 489.

XVII. *The Structure and Properties of Some Ternary Alloys of Manganese, Zinc and Carbon*

By R. G. BUTTERS and H. P. MYERS
Department of Mining and Metallurgy, University of British Columbia, Vancouver 8, Canada*

[Received October 21, 1954]

ABSTRACT

Single phase ternary alloys of Mn, Zn and C have been prepared. These alloys have an ordered structure of the perovskite type and show a spontaneous magnetization at room temperature. The observed magnetic properties are exceptional and indicate a new type of ferrimagnetic behaviour postulated by Néel but hitherto unobserved in experiment.

§ 1. INTRODUCTION

IN the course of work on binary alloys of manganese and aluminium G. R. Piercy noted the occurrence of magnetic properties in certain alloys. It was found that these alloys contained carbon as impurity. Subsequently further work identified the magnetic phase as a ternary alloy of manganese aluminium and carbon possessing a face centred cubic structure, and single phase ternary alloys were obtained over a wide range of composition. Further work by Morgan and Butters (Morgan 1954) showed that magnetic alloys of similar structure could be produced when the aluminium was replaced by indium, tin or zinc.

The magnetic properties of these alloys are being investigated and this paper deals with the results obtained for the manganese zinc carbon system, these alloys having particularly interesting magnetic behaviour.

§ 2. PREPARATION OF THE ALLOYS

The materials used in the preparation of the alloys for this work were manganese of 99·9% purity donated by the Electromanganese Corporation of America, zinc dust of 98% purity† and graphite of spectroscopic grade.

The use of zinc dust was necessary because the alloys were produced by sintering the intimately mixed powdered components. In some cases a master alloy of manganese and carbon was first prepared by induction heating under an argon atmosphere. This alloy was then crushed and mixed with the appropriate amount of zinc dust to produce the mixture for sintering. The mixture of powders was pressed into a clean fused quartz tube and held in place by a quartz plug. The tube was evacuated and sealed, and the sintering achieved by heating to 900°C for 72 hours.

---

* Communicated by Professor W. Sucksmith.

† The major impurity being oxygen.

For alloys containing less than about 15 atomic % zinc this treatment produced hard compact rods of sintered material whose composition was close to the proportions of the materials used.

In alloys with zinc contents greater than 15 atomic % a separation of zinc invariably occurred, the zinc condensing at points where there was free space; this separation of zinc was accompanied by the rejection of some graphite from the sinter. To avoid this loss of zinc by evaporation a lower sintering temperature of 600°c was also used. The zinc loss was thereby reduced but not completely eliminated, and it was not possible to produce an alloy of the desired optimum composition. The sinters were homogeneous but not so hard or compact as those prepared by heating at 900°c.

The alloys decompose in moist air but not so quickly as the isomorphous aluminium alloy which at certain compositions will disintegrate completely in a few hours. They were therefore kept in a dessicator.

Single phase alloys were produced containing varying amounts of zinc; they all showed a spontaneous magnetization at room temperature and were thought to be ferromagnetic. A full investigation, however, revealed strong evidence for their being classed as ferrimagnetic substances.

§ 3. STRUCTURE AND LATTICE PARAMETER

The homogeneity of the alloys was determined by taking Debye-Scherrer diffraction photographs of powders selected from different portions of the sintered mass. Iron radiation and a manganese filter were used. Although the sinters were very hard the material was capable of plastic distortion and sharp diffraction lines were observed only with annealed powders.

Within the range of composition studied, C–20 atomic %, Zn–10–20 atomic %, Mn-70–60 atomic %, the x-ray diffraction patterns showed the structure to be face centred cubic and the presence of extra diffraction lines indicated pronounced ordering. It was apparent that the ordering would be most complete for an optimum atomic composition Mn_3ZnC, and the ordered structure has been determined for an alloy close to this composition by comparing the intensities of the diffraction lines measured by a geiger counter spectrometer, with the calculated intensities expected from an assumed structure.

The diffraction lines typical of a face centred cubic structure were strong and this indicated that it was unlikely that carbon occupied a face centre or cube corner position. The structure assumed was the perovskite type shown in fig. 1. The calculated intensities of the x-ray reflections for this structure are given in table 1, together with the measured intensities. Assumptions were necessary in these calculations because FeKα radiation has a wavelength close to the absorption edge of manganese and the atomic scattering factor is reduced. The decrement in the atomic scattering factor was taken as $-1{\cdot}70$ for Zn, $-4{\cdot}22$ for Mn and 0 for carbon. The temperature factor was omitted.

In comparing the calculated intensities with those observed for a specimen of composition $Mn_{60\cdot9}Zn_{19\cdot1}C_{20}$ reasonably good agreement is found providing a quantitative justification for the assumed perovskite structure. Although diffraction lines of index $\Sigma h^2=1$, 5, 9 and 13 were observed photographically it was not possible to estimate their intensity values with the geiger counter spectrometer.

Table 1

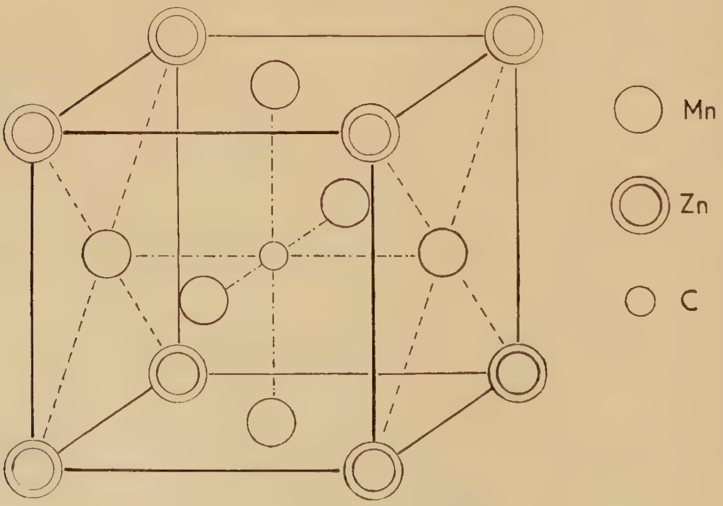
Line index Σh^2	Visual estimate	Intensity Measured	Calculated
1	v. weak	not measurable	0·8
2	weak	11·5	7·8
3	v. strong	100·0	100·0
4	strong	64·9	52·5
5	v. weak	not measurable	0·86
6	weak	3·0	2·63
8	medium	35·4	33·6
9	v. weak	not measurable	0·5
10	v. weak	1·1	0·74
11	strong	45·2	40·5
12	medium	20·0	17·0
13	v. weak	not measurable	0·5
14	weak	7·4	5·1
16	medium	39·0	39·6

The measured and calculated values of the line intensities are based on a value 100 for the strongest line $\Sigma h^2=3$.

Specimen composition $Mn_{60\cdot9}Zn_{19\cdot1}C_{20}$.

Lattice parameter 3·9239 Å.

Fig. 1



Proposed structure of Mn_3ZnC.

The measured lattice parameter of an alloy containing 19·1 atomic % Zn was $3{\cdot}923_9$ Å and extrapolation over a composition range gives the value $3{\cdot}924_9$ Å for the optimum composition Mn_3ZnC.

The ordered structure was extremely stable and occurred when the alloys were slowly cooled or quenched from the sintering temperature.

The phase was stable only for carbon contents close to 20 atomic %, increasing the carbon content beyond this value caused the separation of graphite from the sinter whereas a mixture containing 17·5 atomic % C produced a two-phase alloy on sintering. In all alloys the carbon content was therefore maintained at 20 atomic % and the composition varied by changing the ratio of the manganese to zinc contents. On account of the simple sintering technique the Mn : Zn ratio could not be made less than 3 : 1 due to the difficulties in preventing zinc loss by evaporation. (It is thought, however, that if an improved sintering process made it possible to reduce the Mn : Zn content below 3, useful alloys would not result because the isomorphous Mn, Al, C system shows two phase alloys when the ratio Mn : Al is less than 3, for 20 atomic % carbon.)

A change in composition was produced by increasing the Mn : Zn ratio to values greater than 3 : 1, and several alloys with varying zinc content were prepared but at 8 atomic % Zn instability occurred, the sinter being composed of two coexisting face centred cubic structures with different lattice parameters.

A preliminary determination of the structure of the alloy containing 19·1% zinc at low temperatures was kindly made by Dr. W. B. Pearson of the National Research Council, Ottawa. A Debye–Scherrer x-ray diffraction study was made using the method of Smith and Lonsdale in which liquid nitrogen is passed over the specimen. These experiments showed that at −196°c the original face centred cubic structure of the alloy observed at room temperature is slightly distorted becoming face centred tetragonal with $a = 3{\cdot}921$ Å, and $c/a = 0{\cdot}9947$. Superlattice lines were also present indicating that the tetragonal structure was also ordered.

§ 4. MAGNETIC PROPERTIES

All the alloys studied possessed a spontaneous magnetization at room temperature. The saturation intensities of alloy specimens were measured on a ring balance* similar to that designed by Sucksmith (1939). This apparatus permitted measurements to be made from −196°c to the Curie temperature, the high temperature measurements being made *in vacuo* or a reduced atmosphere of argon. Paramagnetic measurements were made on a similar but more sensitive ring balance (Sucksmith 1929, Sucksmith and Pearce 1938).

Coarsely powdered specimens were found to saturate easily in magnetic fields greater than 7000 oersteds and all ferromagnetic measurements were subsequently made in a field strength of 16 200 oersteds.

Table 2 contains the measured Curie temperature and values of saturation intensity at 20, −45, and −196°c for alloys of different composition. The alloy closest to the optimum composition has the lowest Curie temperature, the latter parameter increasing progressively with increasing Mn : Zn ratio, which corresponds to a decreasing lattice spacing. The graph in fig. 2 shows the variation of Curie temperature with lattice parameter.

---

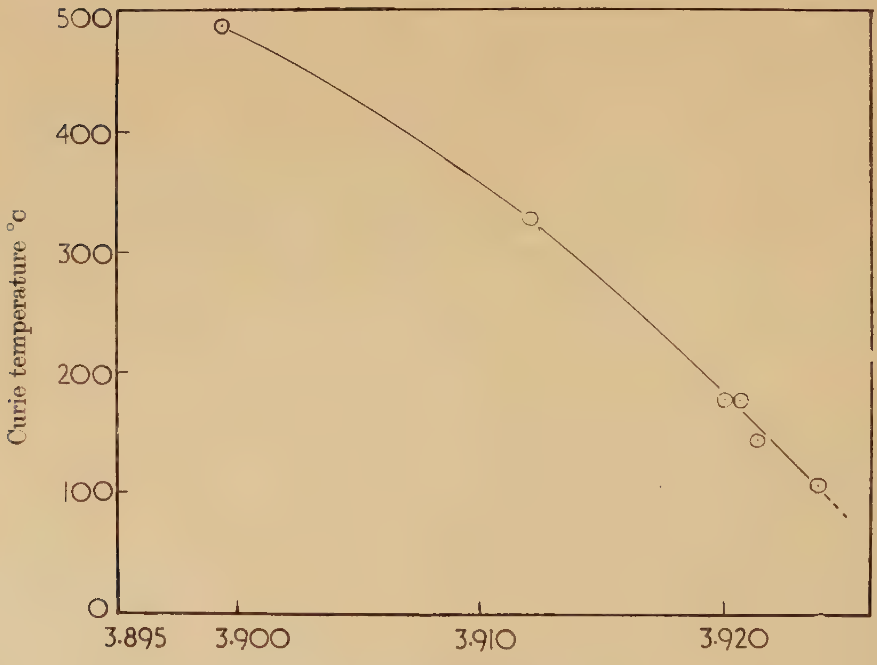
* We are indebted to the Beryllium Corporation of America for the gift of Beryllium copper strip which was used in this apparatus.

Table 2

Composition atom %	Mn : Zn atom ratio	Lattice parameter A	σ_H=16 200 at −196°C	−45°C	20°C	Curie point °C
$Mn_{70\cdot4}Zn_{9\cdot6}C_{20}$	7·32	$3{\cdot}899_4$	61·6	60·4	58·6	488
$Mn_{66\cdot5}Zn_{13\cdot5}C_{20}$	4·94	$3{\cdot}912_4$	82·5	79·0	74·0	325
$Mn_{63\cdot5}Zn_{16\cdot5}C_{20}$	3·86	$3{\cdot}920_7$	83·6	88·6	77·0	177
$Mn_{62\cdot1}Zn_{17\cdot9}C_{20}$	3·48	$3{\cdot}921_4$	77·1	89·0	76·0	143
$Mn_{60\cdot9}Zn_{19\cdot1}C_{20}$	3·18	$3{\cdot}923_9$	71·7	85·5	70·0	107
$Mn_{60}Zn_{20}C_{20}$	3·0	$3{\cdot}924_9$	—	—	—	80

The values quoted for Mn_3ZnC were determined by extrapolation.
Curie temperatures were obtained using σ^2—T graphs.

Fig. 2



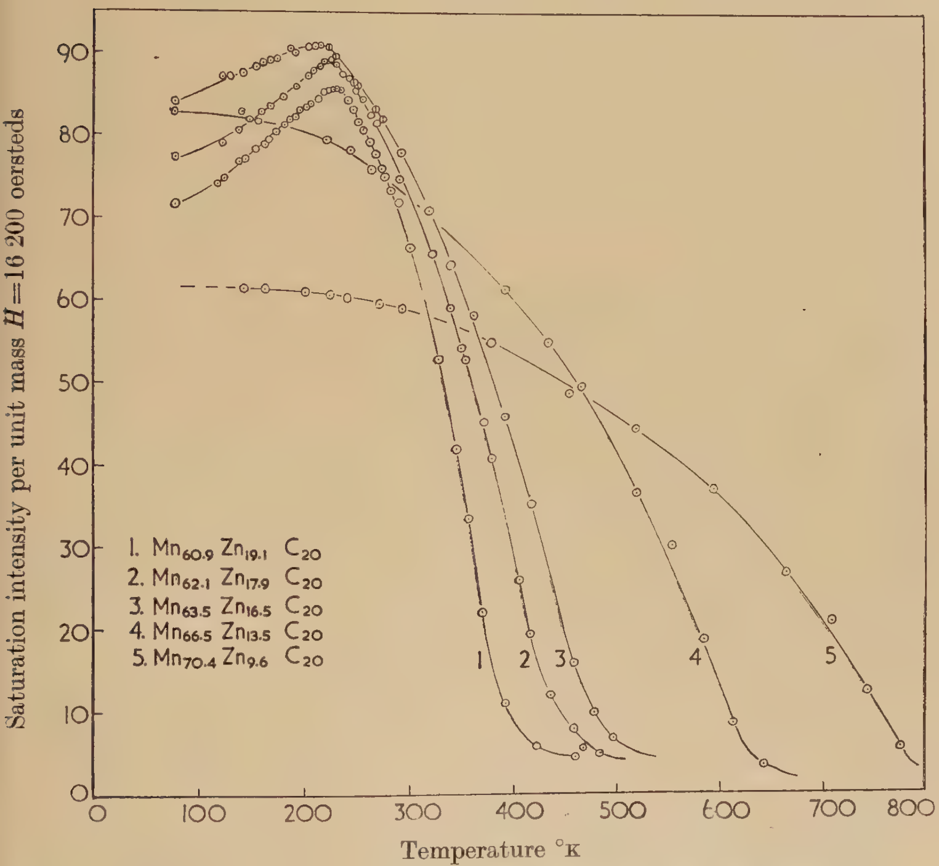
Lattice parameter Å

The variation of Curie temperature with lattice parameter.

In fig. 3 the temperature variation of the saturation intensity is shown. The most interesting feature is the behaviour at low temperatures. For alloys with large zinc content a marked maximum in the magnetization occurs in the region of −40 to −50°C. Above this temperature the magnetization decreases in the usual fashion becoming zero at the Curie point.

Below this temperature region the magnetization decreases but the decrease becomes less the lower the temperature ; the graph is slightly convex towards the temperature axis. Unfortunately on account of the lack of appropriate cryostats we were unable to make measurements at still lower temperatures and extrapolation to 0°K although possible is somewhat uncertain. It is evident however, that the magnetic saturation intensity at 0°K will be only slightly lower than that at −196°C, although measurements at still lower temperatures are needed to establish accurately the manner the magnetization varies as the absolute zero of temperature is approached.

Fig. 3



The variation of the spontaneous magnetization with temperature.

This maximum in magnetization was observed in several alloys of different composition. As the composition was varied so that Mn : Zn atomic ratios became progressively greater than 3 the decrease at low temperatures became less marked until for alloys of about 15 and 10 atomic % Zn the maximum disappeared and the magnetization became constant at low temperatures.

Usually the low temperature measurements were made as the specimen warmed after being immersed in coolant, the temperature being measured by a copper constantan thermocouple; for reasons expressed in a later section measurements were also made as the specimen cooled to low temperatures and we were able to show that the same variation was obtained. It is important to stress the fact that the variation of the magnetization in the region of —45°c was completely reversible with temperature; no thermal hysteresis was observed.

It is interesting to note that, if we extrapolate the low temperature magnetization values and attribute the magnetic moment to the manganese atoms alone, in an alloy of composition Mn_3ZnC the manganese would have an atomic moment of approximately one Bohr magneton per atom at 0°K. The calculated saturation intensity per unit mass of Mn_3ZnC with atomic moment $1\,\mu_B$ for the manganese atoms would be 69 erg/g/oersted.

Paramagnetic measurements above the Curie temperature were made on the alloy having composition nearest to the optimum composition Mn_3ZnC. Unfortunately only a narrow temperature range was available for accurate measurement, for although measurements were made in a reduced atmosphere of argon it was impossible to prevent the loss of zinc by vaporization and an extended range of temperatures could not be used without causing serious error; as it was measurements were made up to 650°c. Invariably measurements taken as the specimen cooled gave smaller susceptibility values than those obtained as the temperature was increased. By comparing values obtained during the heating cycle with those obtained when the specimen was cooled quickly we have attempted to correct for this zinc loss, but it should be emphasized that the values obtained are nevertheless uncertain and can only show the trend of the behaviour at high temperatures. Subsequently resistance measurements made under a full atmosphere of argon showed that zinc distills from the high zinc content alloys at about 400°c and on this account it is thought that accurate susceptibility values will only be obtained by the use of positive pressure apparatus.

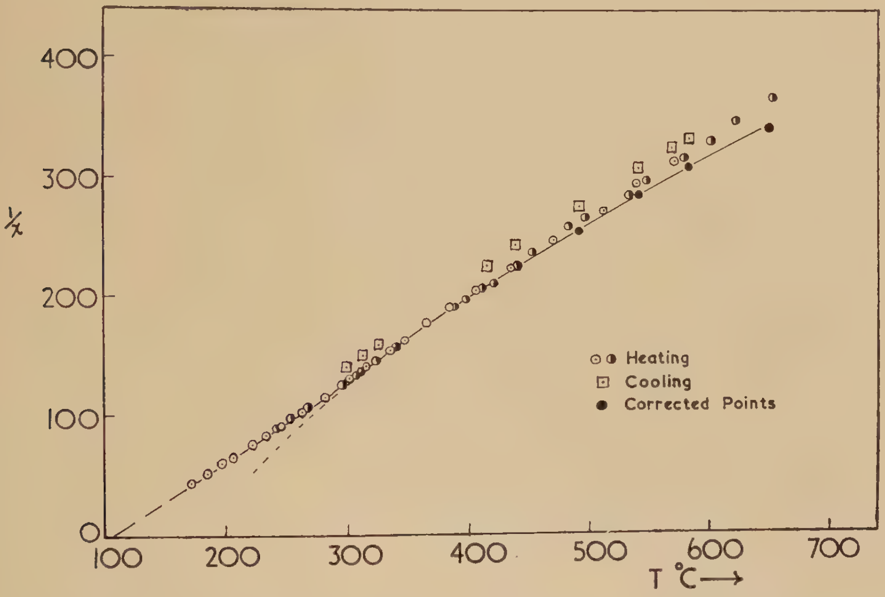
In spite of these difficulties the trend of the variation of $1/\chi$ with T could be determined, and all our measurements showed the $1/\chi - T$ graph to have slight but distinct curvature concave to the temperature axis. This variation is shown in fig. 4.

Electrical Conductivity

Usually the sintered specimens were quite small (about 2 g) but in order to make electrical resistance measurements a sintered rod was prepared. A potentiometric method was used and measurements under one atmosphere of argon were made up to 550°c. The apparent resistivity of the sintered sample was 770 microhm cm at 20°c. The resistance varied linearly with temperature, the substance having an apparent negative temperature coefficient of $1{\cdot}1 \times 10^{-4}$ per °c. Above 400°c the resistance of the specimen began to increase and on cooling this increase in resistance was maintained, although the temperature coefficient of resistance below 400°C was almost unchanged.

On completion of these measurements the copper contacts and the leads not in contact with the specimen were found to be covered with a layer of bright yellow brass formed through distillation of the zinc. The initial composition of the sample for resistance measurements was $Mn_{63\cdot6}Zn_{16\cdot4}C_{20}$, the apparent density and the x-ray density 4·2 and 6·63 g/cm³ respectively.

Fig. 4



The trend of $1/\chi$ against T above the Curie point. $1/\chi$ is referred to a gram molecule of manganese.

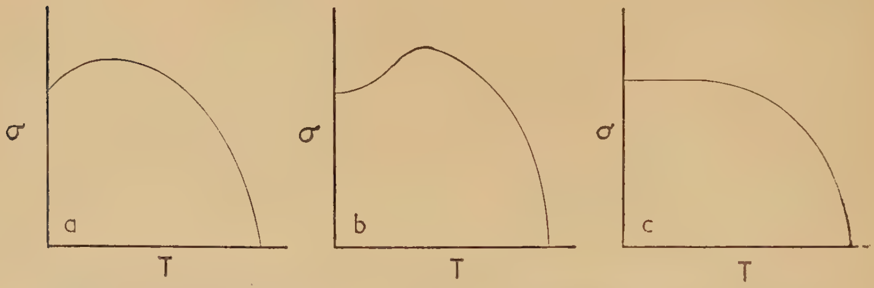
§ 5. DISCUSSION

Although an alloy of exact composition Mn_3ZnC was not studied it is convenient to discuss certain aspects of the observed magnetic behaviour in terms of the structure of the optimum composition ; there are then three manganese atoms per unit cell occupying face centred positions. Extrapolation of the measurements at low temperatures shows that Mn_3ZnC has a magnetic moment of approximately three Bohr magnetons per unit cell at 0°K.

The interesting feature of the magnetic properties is the maximum of the saturation intensity which occurs in the region of −45°C. The observed variation has two possible explanations, arising from structural changes or having a true magnetic origin in ferrimagnetism. The possibility of a structural change is considered first. The magnetization would decrease with temperature below −45°C if a new phase either non magnetic or less magnetic than that stable at higher temperatures was formed below $-45°$C. A phase change at such a low temperature would most probably be a diffusionless transformation and be manifested by an abrupt rather than a continuous change in the magnetization. This type of phase change would also be expected to show pronounced thermal hysteresis on reversing the temperature change and such hysteresis was not detected in the magnetic behaviour.

Although x-ray measurements made at liquid nitrogen temperature showed the structure to be tetragonal we do not consider that the magnetic behaviour is to be attributed to a structural change. The distortion of the face centred cubic cell is very slight ($c/a=0 \cdot 9947$) and we believe that this change of crystal structure may be explained in terms of a ferrimagnetic description of the alloy ; that is the structural change arises from the underlying magnetic properties of the manganese atoms rather than the magnetic properties changing on account of an alteration in structure.

Fig. 5



Three types of ferrimagnetic σ–T characteristics postulated by Néel.

Néel (1948) has shown that the magnetization of a ferrimagnetic substance may vary with temperature in a number of ways dependent upon the distribution and interaction of the magnetic carriers. In particular fig. 5 shows those types of ferrimagnetic behaviour of interest in connection with the present work ; parts (*a*) and (*b*) of fig. 5 both show a maximum in the magnetization and differ only in the manner the magnetization varies as the absolute zero of temperature is approached.

The observed behaviour of Mn_3ZnC is in keeping with that of a ferrimagnetic substance having a magnetization temperature characteristic similar to that of fig. 5 (*b*). For the alloys with Mn : Zn atom ratios of 4·94 and 7·32 the magnetization no longer shows a maximum but becomes constant in the low temperature range studied. This variation is similar to the ferrimagnetic characteristic shown in fig. 5 (*c*).

That the magnetization temperature characteristics of the alloys change from type (*b*) to type (*c*) of fig. 5 as the composition changes is in agreement with the general results of Néel (*loc. cit.*) who has shown that this is one of the trends to be expected when the composition or the intermolecular field constants are changed. It is possible however that an alloy of exact composition Mn_3ZnC may vary in the manner of fig. 5 (*a*) and the type of magnetic characteristic change through type (*b*) to type (*c*) of fig. 5 because the curvature at low temperatures observed in the specimen of composition $Mn_{61}Zn_{19}C_{20}$ (Mn : Zn ratio 3·2) is slight and in this respect measurements at lower temperatures than −196°c are needed to establish accurately the low temperature behaviour.

Ferrimagnetic substances, as shown by Néel and exemplified by the studies on ferrites, have a distinctive paramagnetic behaviour above the Curie temperature, the susceptibility χ is given by

$$1/\chi = T/C + 1/\chi_0 - \sigma/T - \theta, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

C being the Curie constant, and $1/\chi_0$, σ and θ parameters for a given substance. The eqn. (1) represents a hyperbola concave to the temperature axis. Although the measured susceptibility values of the alloy $Mn_{66\cdot 9}Zn_{19\cdot 1}C_{20}$ are uncertain the trend of the variation is in keeping with an equation of type (1), the observed curvature (except in the neighbourhood of the Curie temperature) being concave to the temperature axis and not convex as is usual for a ferromagnetic substance.

If the alloy Mn_3ZnC is ferrimagnetic then the two sublattice model rigorously investigated by Néel should be applicable as it is for the ferrites. For these latter substances, however, their structure allows a non arbitrary division of the magnetic ions to form the sublattices and the state of the ions is known. At room temperature the manganese atoms in Mn_3ZnC occupy equivalent lattice sites and their state of ionization is not known ; the structure at low temperature, however, destroys the equivalence of the manganese atom sites and below a simple model is described which accounts qualitatively for the observed structure and magnetic properties.

On account of the observed magnetic moment per unit cell of Mn_3ZnC it seems justifiable to consider simple ratios. The simplest division is to assume that the manganese atoms on the side centres form one sublattice (A) containing two atoms per unit cell and the end centres form the second sublattice (B) possessing one atom per unit cell of Mn_3ZnC.

If, furthermore, integral atomic moments are assumed for the manganese atoms (as for instance in the Heusler alloys) the observed magnetic moment of Mn_3ZnC can be obtained by the arrangements described in table 3. For ferrimagnetism the magnetization of the sub-lattices A and B are directed oppositely and table 3 gives the states of the ions A and B that would account for the observed moment.

It is unlikely that manganese ions in widely different states would coexist in Mn_3ZnC and choice d is not considered on this account : in addition the number of free electrons available makes the choice c unsuitable unless the carbon atoms contribute electrons to the d shell of the manganese. Furthermore the presence of manganese ions in different states would distort the structural cell from that of the face centred cubic which was accurately observed at room temperature. The most probable state for the manganese ions is therefore that with seven 3d electrons providing each ion with an atomic moment of 3 μ_B.

It is unfortunate that the volatility of the zinc content in these alloys prevents the accurate measurement of the paramagnetic susceptibility over a wide temperature range and thereby precludes the determination of the intermolecular field coefficient n, α and β (see Néel, *loc. cit.*). It is not possible therefore to calculate the variation of the magnetization with temperature as described by Néel.

Table 3

	State of 3d shell ion A	State of 3d shell ion B	Resultant moment per unit cell in Bohr magnetons
a	$3d^6$	$3d^5$	3
b	$3d^7$	$3d^7$	3
c	$3d^8$	$3d^9$	3
d	$3d^9$	$3d^5$	3

Using the ferrimagnetic description outlined above the magnetic properties may be described qualitatively as being similar to those postulated by Néel. The low temperature structure may be explained in terms of the attractive and repulsive forces exerted between ions with oppositely and similarly directed resultant atomic spins (see Zener 1951). These forces will be stronger the lower the temperature, the individual sub-lattices possessing greater magnetization the lower the temperature. If this is the case the result of these forces will be so as to expand the structural cell in the basal plane (taken as the 001 plane in fig. 1) and to contract the cell in the direction normal to this plane; if these forces are strong enough the original face centred cubic cell will be distorted to become tetragonal with axial ratio slightly less than unity. Using this description it is to be expected that the axial ratio would change progressively as the temperature is lowered: x-ray measurements to determine whether this is in fact the case are projected.

With regard to the role of the carbon atom occupying the body centre position, the variation of lattice parameter with zinc content shows the zinc atom to have slightly greater atomic diameter than the manganese. These diameters are estimated to be $2{\cdot}75_9$ Å and $2{\cdot}79_2$ Å for the manganese and zinc atoms respectively. Applying Hägg's empirical rules (see Hume-Rothery and Raynor 1954) for the formation of interstitial phases in alloys of face centred cubic structure it is to be expected that the diameter of the carbon atom should lie between the limits 1·13 and 1·63 Å. The carbon atom must therefore be presumed bonded to its six equidistant manganese neighbours.

The manganese zinc carbon alloys all showed electrical conductivity, a particular alloy of composition $Mn_{63{\cdot}6}Zn_{16{\cdot}4}C_{20}$ in sintered form possessing an apparent resistivity of 770 microhm cm. This value is probably much greater than the true resistivity of the alloy by a factor of the order × 10 on account of the porosity of the sintered mass, it may therefore be said to have electrical resistivity typical of a metallic alloy. The observed negative temperature coefficient of resistance is exceptional but this may be due to the porosity of the specimen causing resistance changes which mask a true metallic behaviour.

If we accept that the manganese ions possess seven 3d electrons, the arrangement may be obtained by assuming that the two 4s electrons of each atom enter the 3d shell when manganese combines to form Mn_3ZnC, requiring the valency electrons of the zinc atoms to be shared by the manganese and zinc atoms providing 0·5s electrons per atom ; this would account for the observed conducting properties.

ACKNOWLEDGMENTS

The authors wish to acknowledge the help received through discussions with their colleagues of the Mining and Metallurgy Department, the University of British Columbia, and the encouragement of Professor F. A. Forward who also provided laboratory facilities. Our special thanks are due to Dr. W. B. Pearson who kindly made the low temperature x-ray measurements.

The work described was carried out with the aid of funds provided by the Defence Research Board of Canada, through grant D.R.B. 281.

Note added in proof.—Recent low temperature x-ray studies made by the authors have shown that the structure of the alloy Mn_3ZnC changes in the temperature region where the magnetization is a maximum. The lattice parameters change in a continuous fashion indicative of a second order transition occurring in this temperature region. A preliminary neutron diffraction investigation kindly made by Dr. B. N. Brockhouse of Atomic Energy of Canada, Ltd. provides further evidence to support this point of view. In addition these neutron diffraction results indicate that at high temperatures the alloy behaves as a ferromagnetic substance but at low temperatures there occurs a gradual ordering of the atomic magnetic moments. It would appear that the second order transition previously mentioned is to be associated with a change from ferromagnetic to ferrimagnetic properties but in order to understand the detailed behaviour of the alloys further x-ray and neutron diffraction studies are needed. This work is in progress and the results will be published at a later date.

REFERENCES

HUME-ROTHERY, W., and RAYNOR, G. V., 1954, *The Structure of Metals and Alloys* (London : Institute of Metals).
MORGAN, E. R., 1954, *J. of Metals*, A.I.M.M.E., **6**, 983.
NÉEL, L., 1948, *Annales de Physique*, **3**, 137.
SUCKSMITH, W., 1939, *Proc. Roy. Soc.* A, **170**, 551.
SUCKSMITH, W., and PEARCE, R. R., 1938, *Proc. Roy. Soc.* A, **167**, 189.
ZENER, C., 1951, *Phys. Rev.*, **81**, 440.

XVIII. *Observations on Negative K-Mesons*

By M. W. Friedlander, D. Keefe and M. G. K. Menon,*
R. H. W. Johnston and C. O'Ceallaigh†
and A. Kernan‡

[Received December 13, 1954]

Summary

Three examples of negative heavy mesons which produce nuclear disintegrations at the end of their ranges in emulsion, are described. Certain characteristics of the parent and capture stars of these mesons are discussed.

§ 1. Introduction

Three examples of nuclear disintegrations produced by slow negative heavy mesons in photographic emulsions have been found in the course of general scanning at Bristol and Dublin. Details of these are given in §§ 2, 3 and 4 respectively. Certain features of interest connected with these disintegrations and the stars from which the K^--mesons were ejected are treated in § 5 in relation to some general characteristics common to events of a similar nature previously reported.

The three events were observed in stacks of Ilford G5 stripped emulsions (flights S30 and S35) flown during the International Sardinian Expedition in Summer 1953.

§ 2. Event K^- (Br 1)

(*a*) *The Primary Particle*

The K-meson has a range of 12·23 mm in five emulsion strips. Mass measurements using standard procedures yield the following values :

(i) By the $\hat{\alpha}R$ method : $m_{K^-}=750\pm 120\ m_e$;
(ii) by the constant sagitta method, using the cell-size scheme of Fay, Gottstein and Hain (1954) : $m_{K^-}=590\pm 130\ m_e$;
(iii) by grain density *v.* range : $m_{K^-}=840\pm 100\ m_e$.

(*b*) *The Secondary Disintegration*

On coming to rest in the emulsion the K-meson produced a characteristic σ-star with three branches, all of which ended in the stack, with

---

* H. H. Wills Physical Laboratory, University of Bristol.

† Institute for Advanced Studies, Dublin.

‡ Physics Department, University College, Dublin (now at University of Rochester, N.Y.).

Communicated by Professor C. F. Powell, F.R.S.

respective ranges $5\,\mu$, $5920\,\mu$ and 17·27 mm. The first of these is too short to be identified; the second is steep and from its appearance seems most likely to be a proton; the third was produced by a proton with an energy at emission of 71·5 MeV. Thus the visible energy in this disintegration is about 110 MeV.

(*c*) *The Parent Star of the K^--meson*

The parent star is of type 3+0n, and of the branches, only that due to the K-meson can be identified unambiguously. One of the other tracks ends in the stack with a range of 3·72 mm and is very steep. The remaining particle had a projected track length of about 0·6 mm per plate and left the stack after traversing 25·7 mm of emulsion. From the rate of change of grain density along the track it appears that the mass of this particle is consistent with that of a proton or deuteron; if a proton, its energy at emission was about 128 MeV. On the other hand, if one assumes that the track was caused by a charged hyperon, the time of flight before the particle left the stack was $1{\cdot}8\times10^{-10}$ sec, during which no decay occurred.

§ 3. EVENT K^- (Br 2)

(*a*) *The Primary Particle*

The track of the K-meson in this event was of rather poor geometry; it had a range of $563\,\mu$ in two emulsion strips and an average dip of about 45°. In spite of this, however, it has been possible by careful gap-length measurements to show that the particle has a mass intermediate between those of the π-meson and proton.

Figure 1 shows a plot of mean gap-length versus range for the K^--meson and for 8 protons and 1 deuteron which have roughly the same dip, lie in the same region of emulsion and were measured in the same layer as was the K-meson. The lengths of all resolved gaps were measured on the appropriate sections of the tracks and the mean gap-length thereby deduced. The measured mean gap-length on the K-meson at a residual range of $449\,\mu$ was $0{\cdot}455\pm0{\cdot}046\,\mu$ whereas that expected at this range in these emulsions for a π-meson is $0{\cdot}655\,\mu$ and for a τ-meson, $0{\cdot}395\,\mu$. It is clear then that the particle involved is almost certainly a K-meson, since the mean gap-width lies more than 4 standard deviations from that for a the corresponding π-meson. The mass value deduced was $630^{+160}_{-120}\,m_e$. This rather low value could perhaps be due to an inelastic collision along its path; in fact there is a sharp deviation of 28° at a distance of $30\,\mu$ from its end.

(*b*) *The Secondary Disintegration*

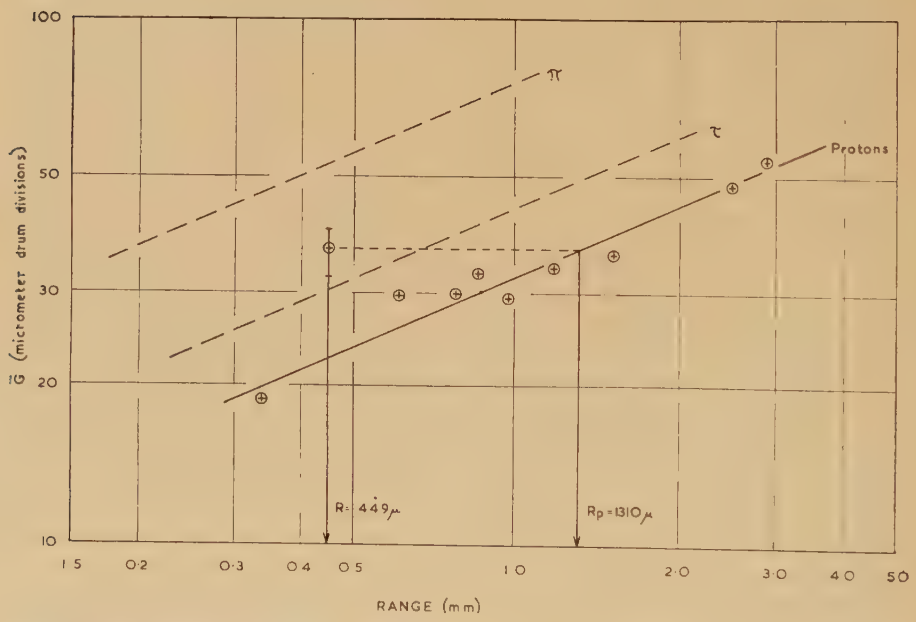
The capture star produced by the K^--meson has four branches, all of which end in the emulsion with respective ranges of $2576\,\mu$, $1180\,\mu$, $356\,\mu$ and $5{\cdot}8\,\mu$. The first has been identified as a proton, the second is probably a deuteron or triton, the third probably a proton or deuteron while the nature of the particle producing the short prong is impossible to establish.

The kinetic energy of the charged fragments lies between 53 MeV and 61 MeV and is probably about 59 MeV.

(*c*) *The Parent Star of the* K^-*-meson*

The parent star is of type 12+0p. In addition to the K-meson, nine of the other eleven particles come to the end of their range in the emulsion and were identified as protons, deuterons or α-particles. Two particles escaped through the edge of the emulsion and were unidentified: the tracks in these cases were steep and grey. No evidence was found for the emission of another unstable particle from the star.

Fig. 1



Plot of the mean gap-length versus range for the tracks of the K^--meson, 8 protons and 1 deuteron.

§4. Event K^- (Db 1)

(*a*) *The Primary Particle*

The K-meson had a range of 19·60 mm in ten plates and the ionization in the neighbourhood of the interaction was consistent with that of a particle of mass ~ 1000 m_e which was at rest or moving with negligible velocity.

Estimates obtained for the mass of the primary particle by different methods were as follows:

(i) Scattering *v.* range (constant sagitta): $m_{K^-} = 780 \pm 100\, m_e$;

(ii) mean gap length *v.* range

from a proton calibration: $m_{K^-} = 950 \pm 60\, m_e$;

from a π-meson calibration: $m_{K^-} = 1100 \pm 100\, m_e$.

(*b*) *The Secondary Disintegration*

On being brought to rest in the emulsion the K-meson produced an insignificant star of three branches, each 2–3μ in length. Thus, no identification of these, the emitted particles, is possible. The visible energy of the star is of the order of 1 MEV.

(*c*) *The Parent Star of the* K^-*-meson*

The star from which the K^--meson emerges is probably of type 4+1p or possible 4+2n. Three black tracks of ranges 40μ, 200μ, and 12μ, and one grey track, that of the K-particle, are arrested in the emulsion ; the two which are at plateau ionization are steep and may be followed out of the emulsion stack.

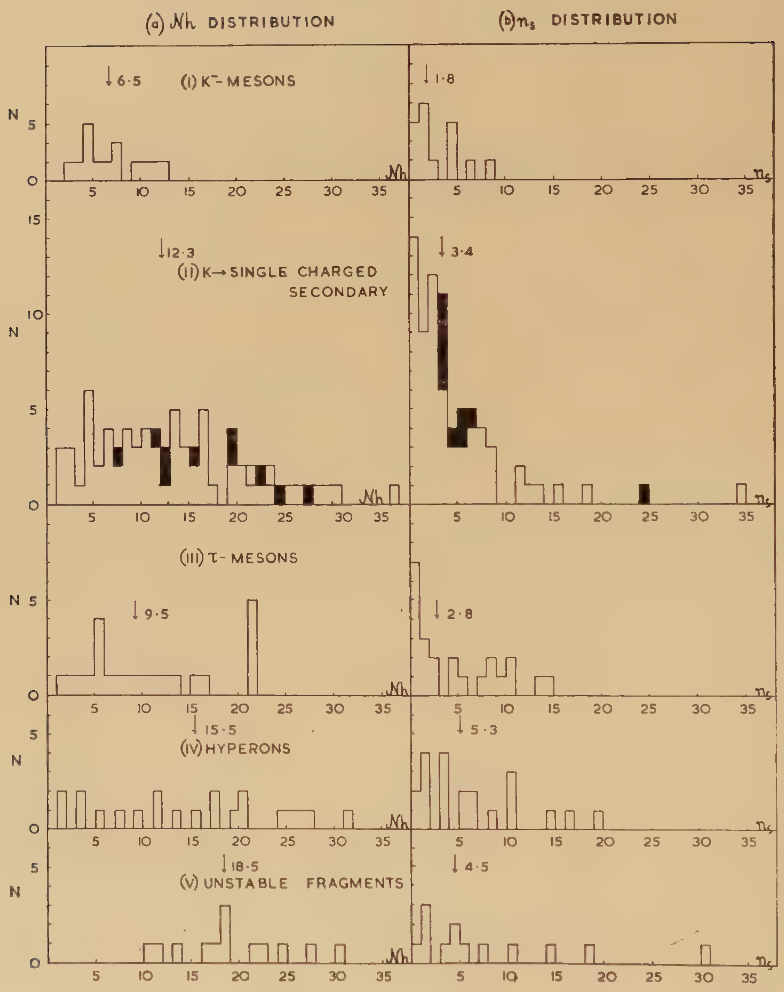
§ 5. DISCUSSION

The total visible kinetic energy in the parent star of event K^- (Br 1) is surprisingly small, being ~ 200 MEV. It appeared worthwhile therefore to examine whether this low energy in the primary star was peculiar to this event or was a general feature characteristic of the parent stars of other K^--mesons. An analysis of similar events hitherto reported shows what may be an interesting, if not easily explicable, property which can be summarized as follows : fig. 2 shows the distribution in N_h and in n_s for the parent stars of the various categories of heavy unstable particle events and from these histograms it would appear that the K^--mesons are, in general, ejected from stars which have small values of N_h, in contrast to those of the other heavy unstable particles. In fact, no well-established K^--meson has yet been observed to originate in a star with $N_h > 12$—a figure less than the median value of N_h for the parent stars of the decaying heavy mesons.

The n_s distributions, on the other hand, show no very significant differences for the various types of particles, though even in this case the median value of n_s for stars from which K^--mesons emerge, is the lowest of all the groups. From a comparison of the histograms of fig. 2 it would appear that the N_h and n_s distributions for the τ-meson and K^--meson origins are not significantly different.

In event K^- (Br 1) the parent star is produced by a neutral particle. If one demands that the K^--meson is produced in a reaction of the type $\pi + N \rightarrow Y + K$, then from charge conservation and the known short lifetime of the neutral π-meson, it follows that here one is dealing with a second generation process. In this connection it is of interest to consider the data summarized in the table, from which it appears that the ratio of charged primaries to neutral primaries for the parent stars of K^--mesons differs from that for the parent stars of the decaying K-mesons. This may perhaps be explicable in terms of the preferential production of negative K-mesons by the charge exchange scattering of neutral heavy mesons. In view of the small statistics involved, however, it must be borne in mind that these apparent systematic differences may be present merely as a result of statistical fluctuations.

Fig. 2



Figures 2 (*a*) and (*b*) show the distributions in N_h and n_s of the parent stars for various types of heavy unstable particles.

The cross hatched squares refer to events observed by Dahanayake *et al.* (1954) in which the particles were found by following out grey tracks from selected stars with $n_s \geqslant 2$, and not by scanning for the meson decays. Besides the bias against $n_s = 0$ or 1, there may also be a possible bias against stars with small N_h owing to scanning loss. The median N_h and n_s values are indicated by vertical arrows.

The data for hyperons and heavy unstable fragments are included here in view of their known association in the production of at least some types of K-mesons.

Primaries	K^-	$\mathrm{K} \rightarrow$ single charged secondary	τ	Y	Unstable fragments
neutral	6	22	10	11	3
charged	5	52	11	10	10

In event K^- (Db 1) the small amount of visible energy associated with the absorption of the meson is striking. From this observation it would appear fairly probable, that in some cases, the capture of the K^--mesons in emulsions will result in nuclear disintegrations with no observable secondary products. Examples of this kind have been reported previously by von Friesen (1954). It is possible to interpret such events, in which almost the entire rest mass energy of the K-meson is carried away by neutral particles, in terms of either of the following processes :

(1) $\mathrm{K}^- + \mathrm{p} \rightarrow \mathrm{Y}^0 + \pi^0 + Q$

(2) $\mathrm{K}^- + \mathrm{p} \rightarrow \mathrm{n} + \mathrm{B}^0 + Q$

where Q=rest mass energy of the K-meson, B^0=a neutral particle and the other symbols have their usual significance. Process (1) is the inverse of the 'associated production mechanism' established by the Brookhaven group. The production of neutral hyperons as a result of the nuclear capture of negative K-mesons has been well illustrated by the observations of Barker (1954) and De Staebler (1954). The kinematics of process (2) have already been considered by Friedlander *et al.* (1954). If $\mathrm{B}^0 = \pi^0$, the energy and momentum considerations for processes (1) and (2) are very similar. If $\mathrm{B}^0 = \nu$ then the process (2) is analogous to the capture of μ^--mesons by all nuclei. There is as yet no direct evidence in favour of process (2).

Acknowledgments

Our thanks are due to Professor C. F. Powell, for the hospitality and facilities of his laboratory. The events were found by Miss Ann Stradling [K^- (Br 1) and K^- (Br 2)] and Miss Moira O'Brien [K^- (Db 1)], to whom we are indebted. Certain of us wish to thank the following bodies : National University of Ireland for a Travelling Studentship (D. K.) ; The Royal Commission for the Exhibition of 1851 for a Senior Studentship (M. G. K. M.) ; and the Board of the School of Cosmic Physics, Dublin, for a scholarship (R. H. W. J.).

References

Barker, K. H., 1954, *Ph. D. Thesis*, University of Manchester.

Dahanayake, C., Francois, P. E., Fujimoto, Y., Iredale, P. I., Waddington, C. J., and Yasin, M., 1954, *Phil. Mag.*, **45**, 855.

De Staebler, H., Jr., 1954, mimeographed preprint.

Fay, H., Gottstein, K., and Hain, K., 1954, *Suppl. Nuovo Cim.*, **11**, 234.

Friedlander, M. W., Harris, G. G., and Menon, M. G. K., 1954, *Proc. Roy. Soc.*, **221**, 394.

Friesen, von S., 1954, *Ark. f. Fysik*, **8**, 305.

XIX. *On the Spin of the θ^0-Meson*

By P. T. MATTHEWS
Department of Mathematical Physics, The University Birmingham
and
ABDUS SALAM
St. John's College, Cambridge †

[Received September 20, 1954]

ABSTRACT

The θ^0-meson may be a scalar or a vector particle. The two possible interactions which account for its decay into two π-mesons are considered. The cross sections (*a*) for the anomalous decay of θ^0 into two π-mesons and a γ-ray, and (*b*) for the production reaction $\pi^- + \mathrm{p} \rightarrow \Lambda^0 + \theta^0$, are calculated, in the hope of finding a criterion to distinguish between the two spin values.

§ 1. POSSIBLE INTERACTIONS

THE existence of a neutral particle θ^0 which decays into a charged π-meson and one other light particle is now firmly established (Butler 1954). On the assumption that the second decay product is also a π-meson, the Q value is 214 MeV.

If the spin of the θ^0-meson is assumed to be less than or equal to one, considerations of angular momentum and parity show that it must be a scalar or a vector particle. The interaction Lagrangian density, giving rise to just the observed decay is thus either

$$L_s(x) = f\theta(x)\,\phi^*(x)\,\phi(x), \qquad (1.1)$$

or

$$L_v{}^1(x) = ig\theta_\mu(x)\left[\phi^*(x)\frac{\partial}{\partial x_\mu}\phi(x) - \frac{\partial\phi^*(x)}{\partial x_\mu}\phi(x)\right], \qquad (1.2)$$

with

$$[\theta_\mu(x), \theta_\nu(x')] = \left(\delta_{\mu\nu} + \frac{1}{M^2}\frac{\partial}{\partial x_\mu}\frac{\partial}{\partial x_\nu'}\right)\Delta(x-x'). \qquad (1.3)$$

Here θ and ϕ are θ^0 and π meson fields and M is the θ^0 mass.

The vector Lagrangian (1.2) is, however, incomplete. In fact if the electromagnetic interaction of charged mesons is also considered, gauge invariance of the theory requires an additional term involving the θ_μ field,

$$L_v{}^2(x) = -2ieg\theta_\mu(x)A_\mu(x)\phi^*(x)\phi(x). \qquad (1.4)$$

We consider now the scalar and the vector interactions for renormalizability. That (1.1) is renormalizable is well-known (Ward 1950). It is remarkable, however, that though the interaction of a neutral vector meson with a Fermi field is renormalizable (Matthews 1949), its direct interaction with a charged boson field is not renormalizable. In this sense interactions (1.2) and (1.4) may be considered as local approximations to the correct interaction.

---

† Communicated by the Authors.

§ 2. CRITERIA FOR DISTINGUISHING SPIN

(*a*) *Anomalous Decays*

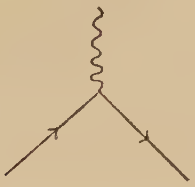
At present a large number (about 20%) of anomalous decays are reported. Just as in the case of π^0 decay (Dalitz 1951), one would expect that a small number (about 1%) of apparently anomalous cases are due to the alternative decay mode

$$\theta^0 \rightarrow \pi^+ + \pi^- + \gamma.$$

On account of the presence of L_v^2 term in L_v, in which this process takes place directly, it was thought that this effect might be more pronounced for the vector than for the scalar theory, and might provide a possible test for the nature of the θ^0. An outline of the calculation is presented below.

The first order Feynman diagram for $\theta^0 \rightarrow \pi^+ + \pi^-$ decay is shown in fig. 1.

Fig. 1



The transition probability is w, where

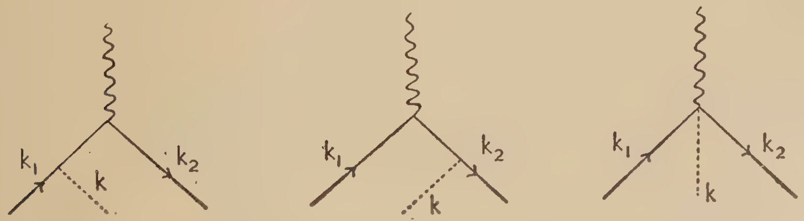
$$w_{2\pi}{}^s = \left(\frac{f^2}{4\pi}\right)\frac{a^{1/2}}{b^2},$$

$$w_{2\pi}{}^v = \left(\frac{g^2}{4\pi}\right)\frac{a^{3/2}}{3b^2}.$$

Here $a = M^2 - 4\mu^2$, $b = 2M$, and μ is π-meson mass.

Figures 2 (*a*) and (*b*) give the lowest order diagrams for scalar 2π, γ decay, while fig. 2 (*c*) gives the additional contribution from L_v^2.

Fig. 2



(*a*) (*b*) (*c*)

The calculation of transition probabilities involves summing over final states of three particles. If $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}$ are the momenta of the two mesons and the photon, respectively, the final integrations involve meson energies w_1 and w_2. Setting $\mathbf{k}_1 + \mathbf{k}_2 = 2\mathbf{k}'$, w_1 for example equals $[(\mathbf{k}' + \mathbf{k}/2)^2 + \mu^2]^{1/2}$.

In the subsequent calculation this has been approximated to by

$$[k'^2+k^2/4+\mu^2]^{1/2}[1+\mathbf{k}\,.\,\mathbf{k}'/(k'^2+k^2/4+\mu^2)]. \quad . \quad . \quad . \quad (2.1)$$

Here k, k' are magnitudes of $\mathbf{k}$ and $\mathbf{k}'$ respectively.

The results are

$$w^{s}_{2\pi,\gamma}=\left(\frac{f^2}{4\pi}\right)\left(\frac{e^2}{4\pi}\right)\frac{2}{3\pi M}\int^{k_{\max}} f(k)\,dk, \quad . \quad . \quad . \quad . \quad (2.2)$$

$$f(k)=\frac{(a-bk)^{3/2}}{k(M-k)(b-k)^2},\; k_{\max}=\frac{a}{b}.$$

It may be noted that (2.2) has an infra-red divergence, so that a lower cut-off for observed photon energies has to be introduced. Taking $k_{\min}=1$ Mev.

$$w^{s}_{2\pi,\gamma}=\left(\frac{f^2}{4\pi}\right)\left(\frac{e^2}{4\pi}\right)\left(\frac{2}{3M\pi}\right)\left(\frac{4(33)^{1/2}}{15}\right) \quad . \quad . \quad . \quad (2.3)$$

and

$$\frac{w^{s}_{2\pi,\gamma}}{w_{2\pi}}\approx(1.58)\left(\frac{e^2}{4\pi}\right)\approx 1.1\%. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.4)$$

For the vector case

$$w^{v}_{2\pi,\gamma}=\sum_{i=1}^{3} w_i{}^{v}=\frac{1}{\pi M}\left(\frac{g^2}{4\pi}\right)\left(\frac{e^2}{4\pi}\right)\sum_i\int^{k_{\max}} f_i\,(k)\,dk.$$

$w_1{}^{v}$ is the contribution from figs. 2 (*a*) and (*b*), $w_3{}^{v}$ from fig. 2 (*c*) and $w_2{}^{v}$ comes from their interference. Here

$$f_1=\frac{1}{9}\left(\frac{k(a-bk)^{3/2}(M-k)}{(b-k)^2}\right)+\frac{M^2(a-bk)^{5/2}}{k(M-k)(b-k)^2},$$

$$f_2=\frac{2}{9}\,\frac{(a-bk)^{3/2}}{(b-k)(M-k)},$$

$$f_3=\frac{1}{6}\,\frac{k(a-bk)^{1/2}}{(M-k)}.$$

$w_1{}^{v}$ is infra-red divergent, and in fact gives the main contribution for $k_{\min}=1$ Mev.

$$w^{v}_{2\pi,\gamma}=\left(\frac{g^2}{4\pi}\right)\left(\frac{e^2}{4\pi}\right)\left(\frac{6.8}{7^2\pi}\right)M, \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.5)$$

and

$$\frac{w^{v}_{2\pi\,\gamma}}{w_{2\pi}{}^{v}}\approx 0.95\left(\frac{e^2}{4\pi}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.6)$$

It thus appears that the ratio of 2π. γ decays (with γ energies greater than 1 Mev) as compared with the total number of 2π decays is slightly greater for the scalar than for the vector case. The difference, however, is not significant. Also, at best, this ratio is of the order of 1%, so that it is not possible to explain most anomalous cases as γ decays.

We obtain a different picture, however, if the emission of only high frequency photons (say in the range 150—165 Mev) is considered. In this case

$$w^{v}\approx\left(\frac{e^2}{4\pi}\right)\left(\frac{g^2}{4\pi}\right)\left(\frac{2}{225}\right)\left(\frac{M}{7^2\pi}\right), \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.7)$$

$$w^{s}\approx\left(\frac{e^2}{4\pi}\right)\left(\frac{g^2}{4\pi}\right)\left(\frac{2}{3M\pi}\right)\left(\frac{1}{8450}\right). \quad . \quad . \quad . \quad . \quad (2.8)$$

This means that events of this type are 10 times as frequent for vector as for scalar interaction, though the probability of 2π, γ decay occurring at all in this frequency range is 10^{-5} times the probability of 2π decay.

§ 3. The Production Reaction

A more promising criterion may be the study of the angular distribution of the reaction

$$\pi^- + \mathrm{p} \rightarrow \Lambda^0 + \theta^0.$$

This reaction is now considered to provide the production mechanism for θ^0 particles. Theoretically, however, one could not set up even a direct scalar renormalizable interaction for this reaction and it must be assumed to proceed via a (virtual) intermediate step $\pi + \mathrm{p} \rightarrow n \rightarrow \Lambda + \theta$. The interaction Hamiltonian for $\Lambda \rightarrow \theta + n$ would be $H^s_{\mathrm{int}} \propto \bar{\psi}_\Lambda \psi_n \theta^0 + \mathrm{c.c.}$ for scalar θ^0 and $\propto \bar{\psi}_\Lambda \gamma_\mu \psi_n \theta_\mu + \mathrm{c.c.}$ for vector θ^0 case. It may be noted that this three (rather than two) field vector interaction is itself unrenormalizable. and is postulated only by analogy. Let $+\mathbf{k}$, $-\mathbf{k}$ be the momenta of the proton and π^- in the centre of mass system, $\mathbf{k}'$, $-\mathbf{k}'$, those of Λ^0 and θ^0, and let angle ϕ be defined by $\mathbf{k} \,.\, \mathbf{k}' = kk' \cos \phi$.

Then $\sigma(\phi) \approx A + B \cos \phi$, for both scalar and vector θ^0.

Here $A_s = 2\, \epsilon_2 \epsilon_3 (k^2 + \epsilon_1 \epsilon_2) - m_2{}^2 (\epsilon_1 \epsilon_3 + m_1 m_3)$,

$$B_s = kk'(2k^2 + m_2{}^2 + 2\epsilon_1 \epsilon_2),$$

$$A_v = 3 m_1 m_2{}^2 m_3 - \epsilon_1 m_2{}^2 \epsilon_3 + 2(k^2 + \epsilon_1 \epsilon_2)\epsilon_2 \epsilon_3 - 2(k'^2 + \epsilon_3 \epsilon_4)\epsilon_1 \epsilon_4 m_2{}^2 / m_4{}^2$$
$$+ 4(k^2 + \epsilon_1 \epsilon_2)(k'^2 + \epsilon_3 \epsilon_4)\epsilon_2 \epsilon_4 / m_4{}^2,$$

$$B_v = -kk'\Big[-2(k^2 + \epsilon_1 \epsilon_2) - m_2{}^2 + 4(k^2 + \epsilon_1 \epsilon_2)(k'^2 + \epsilon_3 \epsilon_4)/m_4{}^2$$
$$+ \frac{2m_2{}^2}{m_4{}^2}(k'^2 + \epsilon_3 \epsilon_4)\Big].$$

Numbers 1, 2, 3, 4 refer to p, π, Λ and θ particles and ϵ stands for the energy. For Brookhaven π^- beam (beam energy 1 500 MEV)

$$\sigma_s(\phi) \propto 2{\cdot}5 + \cos \phi,$$

$$\sigma_v(\phi) \propto 15 - 9 \cos \phi.$$

It must be pointed out that the above cross sections have been calculated using first order perturbation theory. A more accurate calculation would replace γ_5 in pion–nucleon interaction by the vertex operator Λ_5 (Dyson 1949). In view of the unrenormalizability of the direct coupling for vector mesons, one would get into theoretical complications with indirect couplings, if the mesons did turn out to have spin one.

Acknowledgement

We would like to thank Dr. G. Källën for a stimulating correspondence and Mr. R. Shaw for pointing out an error in computation. One of us (P. T. M.) also gratefully acknowledges the hospitality of St. John's College, Cambridge.

REFERENCES

BUTLER, C. C., 1954, *Nuclear Physics Conference, Glasgow.*
DALITZ, R. H., 1951, *Proc. Phys. Soc.* **64,** 667.
DYSON, F. J., 1949, *Phys. Rev.,* **75,** 1736.
MATTHEWS, P. T., 1949, *Phys. Rev.,* **76,** 1254.
STÜCKELBERG, E. C., 1938, *Helv. Phys. Acta,* **11,** 225.
WARD, J. C., 1950, *Phys. Rev.,* **79,** 406.

XX. *The Measurement of Electron Transfer by X-Ray Diffraction Methods*

By P. J. Black
Royal Society Mr. and Mrs. John Jaffe Donation Student
CrystallographicLaboratory, Cavendish Laboratory, Cambridge, England *

[Received November 22, 1954]

Abstract

In the course of structural analyses of aluminium-transition metal compounds, attempts have been made to find evidence for electron transfer effects proposed by Raynor. The results obtained are ambiguous if the intensities have not been measured on an absolute scale. Previous interpretations based on $(F_0 - F_c)$ syntheses must be revised because deficiencies in the *f*-curves used in the calculations have been revealed. It is also shown that the effects of overlap, series-termination and limited counting area can alter appreciably the electron counts on F_0 syntheses. The ultimate possibility of detecting transfer is considered.

§ 1. Introduction

The electron density distribution in a crystal can be derived by Fourier synthesis of x-ray diffraction data. This distribution can be regarded as the summation of a set of individual atoms, and analysis of data in terms of atoms involves the use of *f*-curves, which are theoretical calculations of the scattering due to isolated atoms. As the techniques of measurement and interpretation of diffraction data are improved, details of the continuous distribution are receiving more attention. This paper attempts to examine a particular set of results on electron distribution concerned with a proposed electron transfer in aluminium-rich intermetallic compounds, but the principles involved in the examination are of more general interest.

The background to this work has been reviewed elsewhere (see Raynor 1949, Taylor 1954). For the present purpose, we need only state briefly the problem to which the x-ray techniques have been applied. It has been proposed that in aluminium-rich compounds of the transition metals from chromium to nickel, electrons are 'absorbed' into the 3d shells of transition atoms to an extent which varies from 0·61 electrons for nickel to 4·66 for chromium. Attempts have been made to test this hypothesis by the use of diffraction data. Different and distinct methods of treating the data have been applied by several workers (Douglas 1950, Robinson 1952, 1953, Nicol 1953) and these methods will be classified and distinguished in a review of the results in § 2. Specific details of the interpretation will be examined in subsequent sections.

---

* Communicated by the Author.

§ 2. REVIEW OF THE METHODS USED

(*a*) The first technique uses only observed intensities in the preparation of electron density syntheses (F_0 syntheses). In the maps of electron density, the number of electrons associated with each of the atoms is obtained from the summation of the electron density within an area around the atom centre. Experimental measurements of the intensities used have not been made on an absolute scale, and this means that nothing is known about the existence or magnitude of any continuous distribution of electron density. All that can be measured from the synthesis is the relative weights of atomic peaks above the background. If it is assumed that this background is zero, then the total count over the peaks in a projection can be equated to the total number of electrons given by chemical analysis. Such a method applied to the results for Co_2Al_9 (Douglas 1950) showed that the cobalt atom has 29 electrons instead of its normal complement of 27.

(*b*) In the work of Robinson on β (AlMnSi) (1952) and Nicol on $MnAl_6$ (1953), a different method of scaling F_0 counts was applied. The F_0 values were scaled in each of a series of regions of $\sin\theta/\lambda$ to have the same mean value as a set of structure factors (F_c) calculated with a Viervoll and Ogrim (1949) scattering factor for manganese and a James and Brindley scattering factor for aluminium. When the F_0 values had in this way been placed on an ' absolute ' scale, the F (000) term could be inserted from the total number of electrons per cell and electron counts were read directly from the synthesis. The results indicated an excess of electrons on the manganese atoms and also gave an appreciable (about 30 per cell out of a total of 400) count of electrons in the background between atoms. The interpretation for results on the β (AlMnSi) compound was uncertain because the measured chemical and (ideal) structural compositions differ appreciably, and there is in general no guarantee for these structures that all of the atom sites are always fully occupied.

(*c*) Overlap and series termination may influence counts on F_0 syntheses and (F_0-F_c) syntheses have been used to eliminate their effects. These syntheses show how the electron distribution in the structure deviates from the ideal assumed in the atomic *f*-curves used to obtain the F_c's. The F_0 values have been scaled to F_c values as in (*b*), so that the total count over the difference synthesis must be zero. Because the absorbed electrons are in 3d shells, their scattering is only appreciable at low angles, and an (F_0-F_c) synthesis using only low-angle data should reveal their presence. Syntheses of this type have shown for manganese in β (AlMnSi) an excess of one electron (Robinson 1952) and for manganese in $MnAl_6$, an excess of 1·5–2·0 electrons (Nicol 1953). In another method used by

Robinson (1952), certain low-angle reflections to which the transition atoms make a critical contribution were selected and their observed intensities compared with calculations for different assumptions about electron transfer. This is a variant of the (F_0-F_c) method which employs only particular reflections instead of the synthesis of all of them.

§ 3. THE ATOMIC SCATTERING FACTORS AND (F_0-F_c) SYNTHESES

Methods of type 2 (*b*) and 2 (*c*) depend for their reliability on the *f*-curves employed. Cochran (1953) has suggested a method of checking these using artificial syntheses. This check determines the number of electrons contained within a certain radius in the atom to which the *f*-curve corresponds. This has been applied to the scattering factors for aluminium, silicon and manganese : the procedure of Cochran was followed in detail and the results shown in table 1 correspond to an atom radius of about 3 Å. They indicate that these *f*-curves correspond to atoms deficient in electrons, and in particular, that the Viervoll and Ogrim (interpolated) factors for manganese correspond to a relatively greater deficiency than the factors for aluminium.

Table 1. Check by Artificial Syntheses on Numbers of Electrons Corresponding to Different Atomic *f*-Curves

Atom	*z*		No. of electrons within 3·0 Å of atom centre
Aluminium	13	James and Brindley	12·2
Aluminium 3^+	10	,, ,,	9·6
Silicon	14	,, ,,	13·3
Manganese	25	Viervoll and Ogrim	22·6
3 (3d) electrons	3	,, ,,	3·0
Mn + 3(3d)	28	,, ,,	25·6

In the work of Nicol and of Robinson, the manganese Viervoll and Ogrim (1949) and the aluminium James and Brindley *f*'s were used. Viervoll and Ogrim's results allow an estimate of the effect of extra 3d electrons to be made, and this was used by Robinson.

To show how these results may affect previous interpretations, the results for one synthesis for $MnAl_6$ are discussed. Table 2 shows the peak weights in the asymmetric unit of the (*hk*0) projection, for an F_0 synthesis on an absolute scale for normal atoms (column (*a*)) and for an F_c synthesis using the *f*-curves as in § 2 (*b*) (column (*b*)). The difference is shown in (column (*c*)) but this would not be observed on an (F_0-F_c) synthesis because the coefficients used in (*a*) and (*b*) are first scaled to each other. This can be approximately adjusted by multiplying all figures in (*b*) by a constant factor to give the same total weight as for (*a*) : the results are given in column (*d*). The ($F_0 - F_c$) results of column (*e*) are then obtained. Comparison with Nicol's values shows that his ($F_0 - F_c$) effects are larger than those given by the *f*-curve anomalies, but this calculation leads to the conclusion that the excess of electrons per manganese atom for this synthesis is now only 0·5. It will be shown later, however, that the real excess may be greater than is indicated by the above argument, for it may have been reduced in the operation of scaling F_0 to F_c (cf. §§ 4 (*c*) and (*d*)).

Table 2. Weights of Atomic Peaks in (*hk*0) Projection Syntheses of $MnAl_6$

	(*a*) F_0 synthesis	(*b*) F_c synthesis	(*c*) ($F_0 - F_c$)	(*d*) F_c after scaling to F_0	(*e*) scaled ($F_0 - F_c$)	(*f*) Nicol's ($F_0 - F_c$) result
Mn (2 atoms)	50	45·2	4·8	48·6	1·4	2·4
Al_1 (2 atoms)	26	24·4	1·6	26·2	$\overline{0{\cdot}2}$	$\overline{0{\cdot}7}$
Al_2 (2 atoms)	26	24·4	1·6	26·2	$\overline{0{\cdot}2}$	$\overline{1{\cdot}3}$
Al_3 (1 atom)	13	12·2	0·8	13·1	$\overline{0{\cdot}1}$	$\overline{0{\cdot}2}$

The numbers of the atoms correspond to those used in fig. 3 (*a*) of Nicol (1953).

A similar correction would reduce the effects found by Robinson for β (AlMnSi), both on ($F_0 - F_c$) syntheses and by comparison of low-angle reflections. It must be emphasized that this discussion has proceeded on the assumption that the background of electron density is zero. Because F_0 is scaled to F_c, the difference synthesis can only show abnormality of peak-height ratios, and the interpretation is still ambiguous in the absence of absolute scaling.

§ 4. THE F_0 SYNTHESIS

(*a*) The ambiguity involved in F_0 counts assuming zero background (§ 2 (*a*)) is easily understood. This assumption for Co_2Al_9 led to the result that cobalt has 29 electrons. It could alternatively be assumed that this atom has its normal complement of 27 electrons, that the number of electrons in aluminium is reduced to give the observed ratio of peak weights for cobalt to aluminium, and that the excess of electrons is distributed in a background density. For Co_2Al_9, this latter interpretation would indicate a background of 0·2 electrons per Å³, or, in other terms, would show that in a volume occupied by nine aluminium and two cobalt atoms there are thirteen electrons which are more than 0·7 Å from the atom centres, an alternative explanation which is physically quite reasonable. The distance of 0·7 Å is the radius of the circle over which the count for an atom is made on the projection, and this radius is limited by the overlap of the atomic peaks in the projection.

(*b*) All counts made in this work have been restricted in the above way and even a radius of 0·7 Å usually means that circles from adjacent atoms are overlapping. Consideration of radial charge distribution for Hartree atoms (Hartree 1933) will show that in Cu^+, for example, there are about three electrons lying more than 0·7 Å from the atom centre. About half of these would project outside the counting circle of that radius. A similar estimate for aluminium cannot be made directly: the difference between $U(r)$ curves for Al^{2+} and Al^{3+} (James, Brindley and Wood 1929) shows that the extra electron is almost entirely at more than 0·7 Å from the centre. Al^{2+} has altogether 1·5 electrons outside this radius and it seems reasonable to assume that normal Al would have a greater proportional 'spread' than the transition atom. In that case electron counts on a restricted area would exaggerate the transition metal to aluminium ratio. Thermal vibration or the use of an artificial temperature factor would increase this effect. It is not possible to apply any quantitative estimate because the atoms in the structure are not isolated Hartree atoms.

It is important to interpret correctly the appearance of an F_0 synthesis; this is a map of electron density (ρ) and a very small change in ρ at large distances from an atom centre may correspond to a total change of several electrons in a shell of large radius. The effect will, however, be swamped by the diffraction effects due to the large central peak of ρ; because this peak falls to a first diffraction zero at about 0·3 Å from the atom centre (a typical value for the work considered here), this does not mean that a counting area with even twice this radius will include all significant contributions to ρ.

(*c*) To illustrate the effect of these and other factors, some results obtained in the course of calculations on a projection of the Co_2Al_5 structure will be considered. This structure was solved by Bradley and Cheng (1938) using powder methods. Single-crystal (eye-estimated) intensity data collected by Dr. Newkirk in this laboratory are being used to determine more accurate parameters for the structure. At present, the R-factor for the ($hki0$) zone is 0·095. Counts have been made on F_0 and several F_c syntheses for a projection down the c-axis of the hexagonal unit cell: these calculations are made for a cell based on the orthohexagonal axes. The results are reproduced in table 3, and the positions of the peaks and counting areas are shown in the figure. All of the counts were made on syntheses calculated at 6° intervals; a comparison for one synthesis calculated at 3° intervals showed that peak weight ratios could change by as much as 5% on using the finer grid. The relationship of an atom centre to the grid points of calculated ρ may affect the count, especially if the grid is too coarse. This factor may explain the difference between the cobalt peaks 5 and 6 which are related by hexad symmetry, but appear in different positions relative to the ortho-hexagonal grid.

The F_0 synthesis, scaled by assuming a zero background (method of § 2 (*a*)) would seem to show an excess of electrons on cobalt. An F_c synthesis (F_{c_1}) drawn for Fermi–Thomas cobalt and James and Brindley

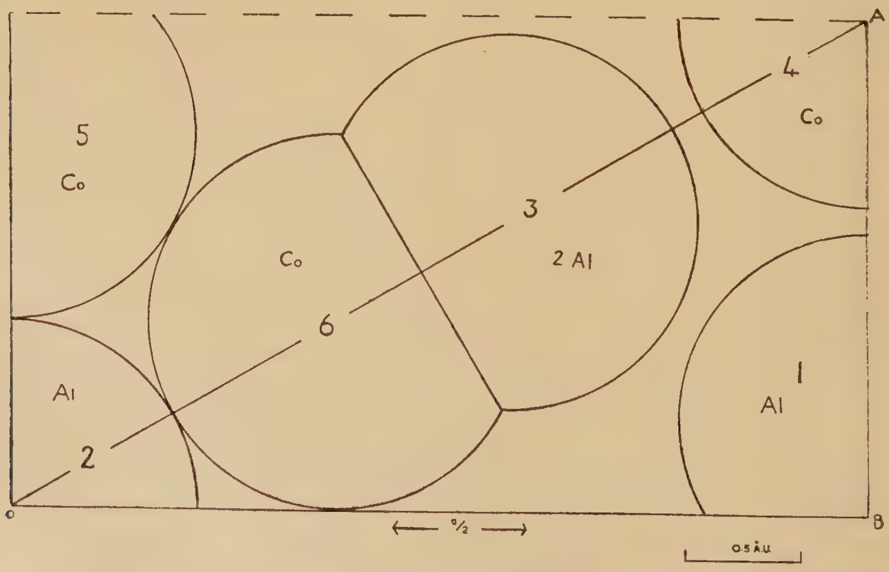
Table 3

		Zero background F_0	Zero background F_{c_1}	Scaled to F_0 zero background F_{c_2}	‘ Absolute scale ’ F_{c_3}
Al	1	9·7	10·2	9·3	12·5
	2	12·3	13·3	12·0	12·8
	3	12·6	13·0	12·8	11·8
Co	4	24·2	27·2	23·7	25·9†
	5	29·4	28·2	29·2	25·3
	6	29·3	27·7	29·6	25·1
*Co/Al mean ratio		2·36	2·20	2·37	2·09

$$* \ (1) \ \frac{\text{Normal Co}}{\text{Normal Al}}=\frac{27}{13}=2{\cdot}08.$$

$$(2) \ \frac{\text{Fermi–Thomas Co}}{\text{James and Brindley Al}}=\frac{26{\cdot}1}{12{\cdot}2}=2{\cdot}14.$$

† Since the counting area for 4 is effectively a complete circle whilst the circles of 5 and 6 are cut off by chords, a change in the background adds more to peak 4 than to the others.



c-axis projection for Co_2Al_5, showing positions of counting areas for peaks. Numbers correspond to table 3. OAB is the asymmetric unit.

aluminium atoms, without a temperature factor, shows individual variations between peaks of the same kind of atom comparable to those of the F_0 synthesis and gives a mean Co : Al ratio a little higher than the f-curves would predict. If the coefficients of this synthesis are scaled to those of the F_0 synthesis, then the synthesis F_{c_2} is obtained. This scaling will include the effect of temperature factor, and low angle variations due to $\alpha_1-\alpha_2$ resolution and extinction (neither of the latter can be effectively corrected). No F (000) term is inserted and the peak weights are obtained by assuming a zero background as in F_{c_1}. The results reproduce the abnormal Co : Al ratio, and the application of such scaling factors can thus make a critical difference to this type of result. The individual variations between peaks however, do show that the combined effects of limited counting area, of overlap and of series termination cannot be generally predicted, and that each synthesis will have its own features, so that results on F_0 syntheses cannot be assessed without careful and individual investigation.

(*d*) F_{c_3} is the same synthesis as F_{c_1}, with an appropriate F (000) term (i.e. corresponding to the number of electrons per cell) inserted. Electron density values are then read directly off the synthesis and there is no assumption about zero background. This is the method of § 2 (*b*). There is in fact an appreciable background of electron density between the atoms in this synthesis and the ratio of peak weights is also changed as they are reckoned from a different background. The existence of this background on an F_c synthesis is explained because the F (000) term corresponds to normal atoms whilst the f-curves correspond to deficient atoms, or alternatively, because the f-curves correspond to atoms whose electron density spreads appreciably beyond the counting circle.

This method of scaling electron counts was used by Robinson (1952) and Nicol (1953) for F_0 syntheses, after the F_0 coefficients had first been scaled to F_c coefficients, so imposing on the F_0 values a mean angular variation corresponding to that of the F_c values. In both cases an appreciable number of electrons was found in the background, but it can be seen by comparison with the previous paragraph that this background is an artefact arising from the method of scaling. The introduction of this background would distort the ratio of peak heights. The scaling to F_c coefficients may give an approximately correct angular variation, but it cannot indicate the correct F (000) term—i.e. it can never show up a background density on an F_0-F_c synthesis.

§ 5. The Detection of Electron Transfer

Bijvoet and Lonsdale (1953) have questioned the possibility of finding electrons in outer atomic shells. They suggest that shells of large radius only affect spectra at low angles where no diffraction spectra exist. In a discussion of the results of previous workers for LiH they criticized the extrapolation of measured f-curves to zero angle, and showed that the difference between the scattering powers of the normal and ionized atoms at the lowest observable angle was negligible. Their conclusion was that no evidence for or against ionization in LiH could be obtained by this means.

The extrapolation of f-curves to zero angle is not justified and the only possibility of detecting any ionic effect lies in the direct measurement of low angle spectra. For the LiH case the ratio F_{111}/F_{200} (where $F_{111}=f_{\mathrm{Li}}-f_{\mathrm{H}}$ and $F_{200}=f_{\mathrm{Li}}+f_{\mathrm{H}}$) is 0·65 for normal Li and H, and has been measured independently by three workers as 0·50, 0·50 and 0·55. There is a significant difference, which is in the right direction for ionization, but it is not possible to say more unless one has correct f-curves for atoms in this structure. In considering the difference in scattering effects between H and H^-, Lonsdale and Bijvoet use an f-curve for H^- which is valid for a free ion but a calculation of its radial electron distribution shows that the second electron is distributed at 0·8 to 2·0 Å from the atom, whilst the interatomic Li–H distance is only 2·1 Å. Waller and Lundquist (1953) have calculated the scattering in this crystal by correcting for the effect of overlapping between neighbouring atoms, and assuming complete ionization, obtained good agreement with observation for their calculation of the F_{111}/F_{200} ratio. They do not indicate the extent to which this agreement depends on the degree of ionization assumed.

One can assert that a complete set of Fourier terms will give a complete representation of the electron density in a structure—there is no question of low-angle ‘ cut-off ’ in the data—but it may not be possible to assign the (small) electron density between atoms to particular shells or particular atoms. If the F_{000} term is not measured, the critical factor is the order of magnitude of the effect of any electron-transfer on the lowest angle spectra. Consider a structure with two different types of atom, separated by a distance d'. If an electron is transferred it should lie in or within a shell of radius $d'/2$, and the f-curve of this shell is given by

$$f=\int_0^\infty U(r)\,\frac{\sin \mu r}{\mu r}\,dr$$

where $\mu=4\pi \sin\theta/\lambda$. For a thin shell, this becomes

$$f=\frac{\sin\,(2\pi \sin\theta/\lambda)\,d'}{(2\pi \sin\theta/\lambda)\,d'}\,.$$

If d is the longest possible inter-planar spacing then the lowest angle spectrum occurs at $\sin\theta=\lambda/2d$, and the contribution of this electron is

$$f_d=\frac{\sin\,(\pi d'/d)}{\pi d'/d}\,.$$

For an NaCl type structure, $f_d=0{\cdot}15$ (the least favourable case). For the complex aluminium-transition metal structures, there are spectra from planes of spacing 6 Å to 12 Å, whilst $d'=2{\cdot}5$ Å. Then $f_d=0{\cdot}76$ to 0·93, and the effects should be detectable especially if the transfer involves two or three electrons. The complexity of the structure is, in this respect, an advantage.

§ 6. CONCLUSION

It is concluded that the existence of electron transfer effects in intermetallic compounds has not been reliably established. It appears (see e.g. § 3) that the effect must be smaller than the transfer required by Raynor's theory. If one can obtain accurate atomic positions for a fairly complex structure and place the intensities by experiment on an absolute scale, then phenomena involving two or three electrons should be detected. The best procedure would be to derive empirical *f*-curves for the two types of atom from the data, and then to calculate the numbers of electrons to which these curves correspond. There is no advantage in using calculated *f*-curves unless these correspond to some significant theoretical model. It may be useful for example to use an *f*-curve for the core (up to and including the 3p shell) of a transition atom which can be calculated theoretically and which one would not expect to be influenced by the structural environment, and use this in $(F_0 - F_c)$ syntheses to show only the 3d electrons.

I am grateful to Professor N. F. Mott and Dr. W. H. Taylor for provision of facilities and for their interest in the progress of this work. I wish to thank the Council of the Royal Society for financial support. I am indebted to Dr. J. B. Newkirk for permission to use his measurements on Co_2Al_5, and to Dr. Cochran and other colleagues for helpful discussion and criticism.

REFERENCES

BIJVOET, J. M., and LONSDALE, K., 1953, *Phil. Mag.*, **44,** 204.
BRADLEY, A. J., and CHENG, C. S., 1938, *Z. f. Krist.*, **99,** 480.
COCHRAN, W., 1953, *Acta Cryst.*, **6,** 812.
DOUGLAS, A. M. B., 1950, *Acta Cryst.*, **3,** 19.
HARTREE, D. R., 1933, *Proc. Roy. Soc.* A, **141,** 282.
JAMES, R. W., BRINDLEY, G. W., and WOOD, R. G., 1929, *Proc. Roy. Soc.* A, **125,** 401.
NICOL, A. D. I., 1953, *Acta Cryst.*, **6,** 285.
RAYNOR, G. V., 1949, *Progress in Metal Physics*, **1** (London : Butterworths).
ROBINSON, K., 1952, *Acta Cryst.*, **5,** 397 ; 1953, *Ibid.*, **6,** 854.
TAYLOR, W. H., 1954, *Acta Met.*, **2,** 684.
VIERVOLL, H., and OGRIM, O., 1949, *Acta Cryst.*, **2,** 277.
WALLER, I., and LUNDQUIST, S. O., 1953, *Arkiv f. Fysik*, **7,** 121.

XXI. *Elastic Scattering of High Energy Electrons by Mercury*

By G. E. Brown
Department of Mathematical Physics, University of Birmingham

and

L. R. B. Elton
Wheatstone Department of Physics, King's College, London*

[Received October 19, 1954]

Abstract

We have attempted to fit experimental results for the scattering of electrons by heavy nuclei at 126, 154 and 183 MeV with curves obtained from various assumed nuclear models by way of a phase shift analysis carried out for the case of mercury. At these energies only nuclear scattering is important and the nuclear model completely determines the scattering. Charge distributions which reproduce the experimental curves either are homogeneous or have a slight dip near the centre of the nucleus and have their main decrease on the surface of the nucleus over a transition range of between 1·7 and $3{\cdot}2\times10^{-13}$ cm. The root mean square radius of the distribution is determined to be approximately $R=1{\cdot}2\times10^{-13}\times A^{1/3}$ cm, to a relative accuracy of about 5%.

§ 1. Introduction

We shall report on phase-shift calculations of coherent Coulomb scattering of electrons by mercury. We consider only scattering from the nuclear charge distribution. Others of the effects usually considered have been discussed previously for this energy range (Schiff 1953) and seem to give only small corrections to the Coulomb scattering. At the energies considered, 125 and 180 MeV, the charge distribution of a mercury nucleus has a radius larger than the electron wavelength, and the scattering can be used to determine the size and shape of the charge distribution.

We shall apply our analysis to the experimental data of Hofstadter, Hahn, Knudsen and McIntyre (1954), and shall show that the scattering is sensitive to the range of the charge distribution and to certain general features, especially the length of transition region in which the charge goes from its average value inside the nucleus to zero.

A survey of nuclear models has been carried out by Yennie, Ravenhall and Wilson (1954), Ravenhall and Yennie (1954) and points to the general conclusions we draw here. We have extended the accuracy of previous calculations by Brenner, Brown and Elton (1954) which were inaccurate at large angles as noted in that paper. The general agreement of present calculations is very satisfying considering the large number of phase shifts and high accuracy required.

---

* Communicated by Professor R. E. Peierls, F.R.S.

§ 2. PROCEDURE

The procedure followed is similar to that outlined by Brenner, Brown and Elton (1954), which will be referred to as I, with some modifications which make the calculation more convenient. The notation of I will be employed.

We are concerned with solving the pair of equations*

$$\frac{dS}{dr}+\frac{k}{r}S-(1-V)R=0, \qquad \frac{dR}{dr}-\frac{k}{r}R+(1-V)S=0, \quad . \quad . \quad . \quad (1)$$

where k is Dirac's quantum number, i.e. $k=j+\frac{1}{2}$ if $j=l-\frac{1}{2}$ and $k=-j-\frac{1}{2}$ if $j=l+\frac{1}{2}$. Here j and l have their usual meanings. The solutions become asymptotically

$$\left.\begin{aligned} R&\sim\cos\left(r+Z\alpha\log 2r-\tfrac{1}{2}k\pi+\chi_k\right),\\ S&\sim\sin\left(r+Z\alpha\log 2r-\tfrac{1}{2}k\pi+\chi_k\right). \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The scattering cross section can be expressed in terms of χ_k (Mott and Massey 1949, p. 76). We first define

$$f(\theta)=\frac{1}{2i}\sum_{k=1}^{\infty} k\exp(2i\chi_k)\,[P_k(\cos\theta)+P_{k-1}(\cos\theta)] \quad . \quad . \quad . \quad . \quad (3)$$

where we have reduced the expression of Mott and Massey to that for high energies ; i.e. we have set equal the two phase shifts corresponding to the same j. Then the cross section is given by

$$\frac{d\sigma}{d\Omega}=|f|^2\sec^2\frac{\theta}{2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

We split χ_k into $\eta_k+\delta_k$ where η_k is the Coulomb phase shift for a point charge and δ_k is the 'anomalous phase shift', which is different from zero for only the first few values of k. As explained in I, we express $f(\theta)$ as $f_P(\theta)+[f(\theta)-f_P(\theta)]$ where f_P is the amplitude which would result from a point charge and $f-f_P$ involves only a few terms in a summation in k. We can easily determine δ_k from a knowledge of R/S of (1) and of the regular and irregular point charge functions at some point beyond the nucleus. We have found a more convenient form than that given in I for the asymptotic series for the regular and irregular Coulomb functions which we denote by R_R, S_R and R_I, S_I. This is

$$\left.\begin{aligned} S_R&=(X+U)\sin\sigma_k+(Y+V)\cos\sigma_k,\\ R_R&=(X-U)\cos\sigma_k-(Y-V)\sin\sigma_k,\\ S_I&=(X+U)\cos\sigma_k-(Y+V)\sin\sigma_k,\\ R_I&=-(X-U)\sin\sigma_k-(Y-V)\cos\sigma_k, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

---

* We choose units such that $\hbar=c=E=1$, and neglect the electron mass in comparison with the electron energy. It has been shown in I that the errors made in this neglect are of order $(m/E)^2$.

where $$\sigma_k = r + Z\alpha \log 2r - \tfrac{1}{2}k\pi + \eta_k \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.1)$$

and we can define X, Y, U, V by a power series expansion,

$$\left.\begin{aligned} X+iY &= A_0+iB_0+\frac{i}{2kr}(A_1+iB_1)+\frac{i^2}{2!(2kr)^2}(A_2+iB_2)+\ldots \\ k(U+iV) &= \frac{i}{2kr}(A_1+iB_1)+\frac{i^2}{(2kr)^2}(A_2+iB_2)+\frac{i^3}{2!(2kr)^3}(A_3+iB_3)+\ldots \end{aligned}\right\} \quad . \quad . \quad . \quad (5.2)$$

where

$$\left.\begin{aligned} &A_0=1,\, B_0=0\,;\; A_1=k^2,\, B_1=0\,;\; A_2=(k^2-1^2)\,k^2, \\ &B_2=2Z\alpha\, k^2\,;\; \ldots\ldots\; A_{s+1}=(k^2-s^2)\,A_s-2s\,Z\alpha\,B_s, \\ &B_{s+1}=2s\,Z\alpha\,A_s+(k^2-s^2)\,B_s. \end{aligned}\right\} \quad . \quad . \quad . \quad (5.3)$$

We have evaluated R_R, S_R, R_I and S_I to five decimal places for $k=1\ldots10$ and for $r=12{\cdot}5$ and 14 for each k. We should be glad to furnish the numerical values of these functions upon request. To obtain five-place accuracy, we have had to evaluate only $\sim k$ terms in the series for the kth point charge functions, since the term (k^2-s^2) occurring in the recurrence formula becomes small for $k\sim s$.

We have evaluated the point charge amplitudes to five decimal places by the method of Yennie *et al.* (1954) and consequently have improved greatly the accuracy of the calculations reported in I, where we had employed the amplitudes given by Feshbach (1952). We give a summary of our point charge amplitudes.*

Point Charge Amplitudes

	35°	40°	45°
Real part of f_P	3·069	2·655	2·2675
Imaginary part of f_P	—2·095	—1·211	—0·6488

50°	55°	60°	65°	70°	75°	80°
1·9261	1·6320	1·3811	1·1684	0·9881	0·8354	0·7058
—0·2870	—0·0538	0·0951	0·1873	0·2413	0·2691	0·2789

85°	90°	95°	100°	105°	110°	115°
0·5956	0·5017	0·42155(5)	0·35302	0·29437(5)	0·24415	0·20114
0·2766	0·2659	0·24981	0·23029	0·20882	0·18644	0·16394

120°	125°	130°	135°	140°	145°	150°
0·16430	0·13278	0·10589	0·08303	0·06372(5)	0·04755	0·03417
0·14189	0·12071	0·10075	0·08224	0·06537	0·05028(5)	0·03707

These amplitudes agree with those of Feshbach to within one unit in the last place given by him at the angles at which they can be compared, except at the smallest angle, 45°. The method used by us does not give accurate results at small angles, but we expect the accuracy at large angles ($\geqslant 90°$) to be to the number of decimals quoted.

---

* The $G(\theta)$ tabulated by Feshbach (1953) is the same as our $f_P(\theta)$.

For the calculations at 125 MeV we have had to evaluate 7–8 anomalous phase shifts and for the calculations at 180 MeV we have employed 10 phase shifts. These phase shifts have been calculated to five decimal places.* Consequently, the theoretical curves are accurately defined as shown in figs. 1–7. High accuracy is important because 2 to 3 significant figures in the amplitudes are lost through cancellation at large angles.

§ 3. RESULTS

We show in figs. 1–7 results from nuclear models with more or less the correct range to reproduce the experimental results. The range of the distribution is defined in terms of the square (homogeneous) distribution which would reproduce the same low energy effects (Feshbach 1952) and this range, in turn, is used to define a nuclear radius parameter r_0 in the following way :

$$\left.\begin{aligned} R_s^2 &= \frac{5}{3}\cdot\frac{4\pi}{Ze}\int_0^\infty \rho(r)\, r^4\, dr, \\ R_s &= r_0 \times A^{1/3} \times 10^{-13}\ \text{cm}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

(In this section c.g.s. units will be employed.)
We have considered charge distributions of the following form :

(*a*) Homogeneous with Gaussian type cut off

$$\rho = K\,\frac{1}{1+\exp\left[(r^2-\alpha^2)/\beta^2\right]}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

(*b*) Wine bottle

$$\rho = K\,\frac{1+r^2/\alpha^2}{1+\exp\left[(r^2-\alpha^2)/\beta^2\right]}.$$

The quantity β^2 is related to the length of transition region over which the charge decreases from an average value to zero. In (*a*), for example, the charge goes from $0{\cdot}88K$ to $0{\cdot}12K$ as r^2 goes from $\alpha^2-2\beta^2$ to $\alpha^2+2\beta^2$; from this we can define a transition region of length $(\alpha^2+2\beta^2)^{1/2}-(\alpha^2-2\beta^2)^{1/2}$ which is approximately equal to $2\beta^2/\alpha$ for $\beta \ll \alpha$. We can relate α to R_s by (6) and can relate K to the total charge Ze and α by requiring the volume integral of the charge density to be equal to Ze. We have chosen $R_s = 1{\cdot}2 \times A^{1/3} \times 10^{-13}$ cm for all distributions calculated here. Thus, all distributions have the same second and fourth moments of the charge density.

In fig. 1 the results of various curves of type (*a*) are shown, for scattering at 125 MeV. The experimental points are for scattering at 126 MeV from Hofstadter *et al.* (1954). The experiments were carried out for gold and for lead, and gave identical results, within experimental error, for these two elements. We shall therefore use them without correction for mercury.

---

* As a check the fitting of the functions integrated out from the origin to the asymptotic Coulomb functions has been performed at two different places. The phase shifts thus obtained differ from each other by only a few places in the last figure and should therefore be correct to the same accuracy.

Absolute measurements have not been made and consequently the experimental points may be moved up or down by a common distance on a logarithmic plot. It is seen, however, from the large angle scattering that by suitable normalization of the experimental points, curves from charge distributions with transition parameter $\beta^2 \simeq 3\alpha^2/16$ to $\alpha^2/4$ fit the experiments reasonably at this energy. (We shall see that somewhat shorter transition regions fit the 154 and 183 Mev data better.)

Figure 2 indicates the effect of a wine bottle distribution. The transition region in this case is similar (but not exactly equal, if we use the above definition) to that of the curve in fig. 1 with $\beta^2 = \alpha^2/4$. The effect of scooping charge from the centre and putting it farther out is to raise the scattering at intermediate angles, 60° to 90°. Although curves of the type (*a*) fit reasonably the experimental points within the limits of error of these, the general trend is easier to reproduce with the type (*b*) curve.

We can, in fact, see what the effects of different modifications in the charge distribution are. (In all such changes we keep the 2nd and 4th moments constant.) Heaping the charge up at the origin, as in an exponential or Gaussian distribution, raises the scattering at small angles in the region 30–70°, as portrayed in fig. 2. Raising the charge at a distance from the origin, as in the wine bottle, raises the scattering at intermediate angles, 60 to 90°. Increasing the transition range decreases the scattering slightly at intermediate angles and markedly at large angles, $\geqslant 100°$. A comparison with the results of Yennie *et al.* (1954) shows that these remarks are borne out by their work as well.

It is necessary to calculate cross sections only at 125 Mev, as long as the points are fitted to sufficiently wide angles because theoretical curves extrapolate with energy in a known way, namely, the ratios of the differential cross sections to the point scattering cross section is the same for different energies E and E' at corresponding angles, or

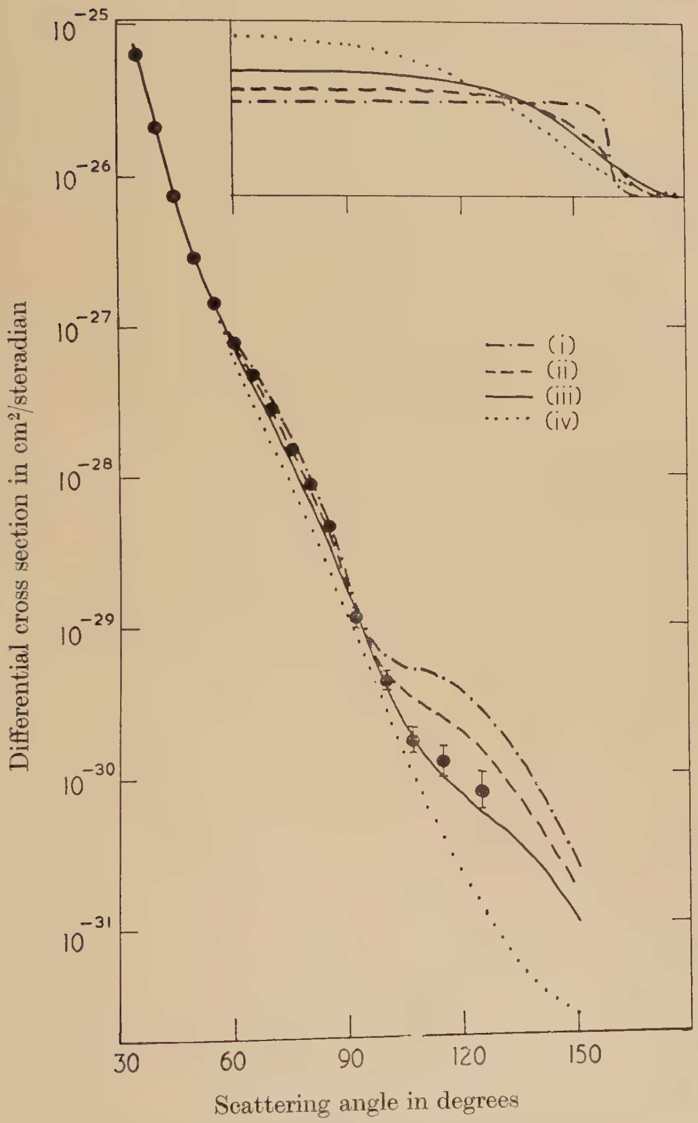
$$\frac{d\sigma(E,\theta)}{d\Omega}\Big/\frac{d\sigma(E,\theta)_P}{d\Omega}=\frac{d\sigma(E',\theta')}{d\Omega'}\Big/\frac{d\sigma(E',\theta')_P{}'}{d\Omega} \quad . \quad . \quad . \quad . \quad (8)$$

where E, E', and θ' are related by

$$(ER)^{0\cdot80}\sin\tfrac{1}{2}\theta=(E'R)^{0\cdot80}\sin\tfrac{1}{2}\theta'. \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

This semi-empirical extrapolation rule was suggested by Dr. McIntyre of Stanford and applied by Yennie *et al.* (1954). To illustrate the utility of this rule, we exhibit in fig. 3 theoretical curves at 180 Mev and curves for the same energy obtained by extrapolating the 125 Mev curve. In fig. 4 the experimental points at 154 and 183 Mev are shown and the curve obtained by extrapolating the curve passing through the centres of the 126 Mev points. (The differences between theoretical curves for 125 and 126 Mev and those between curves for 180 and 183 Mev are quite negligible. In the following figures 5–7 we have, however, performed the small extrapolation from 125 to 126 Mev and from 180 to 183 Mev to make a direct comparison with experiment.) The extrapolated experimental curve does not pass through all of the 183 Mev points, but the experimental

Fig. 1



Differential cross sections for scattering at 125 MeV from nuclear models of the type

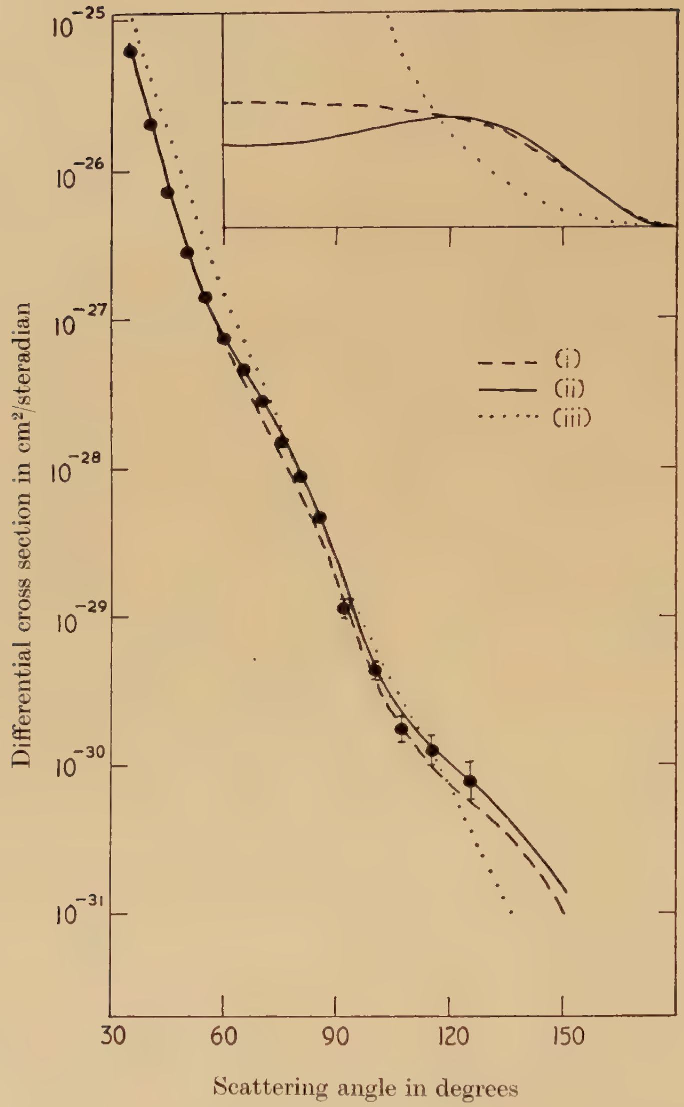
$$\rho = K \frac{1}{1+\exp[(r^2-\alpha^2)/\beta^2]}$$

where the following values of K and α are associated with the given β^2 :

(i) $\beta^2=\alpha^2/16$, $\alpha=4{\cdot}40$, $K=0{\cdot}00159$;
(ii) $\beta^2=\alpha^2/8$, $\alpha=4{\cdot}31$, $K=0{\cdot}00171$;
(iii) $\beta^2=\alpha^2/4$, $\alpha=3{\cdot}91$, $K=0{\cdot}00216$;
(iv) $\beta^2=\alpha^2/2$, $\alpha=3{\cdot}27$, $K=0{\cdot}00300$.

(Lengths are measured in units of $\hbar c/E$, energies in units of E.) All distributions have nuclear parameter $r_0=1{\cdot}2$ (in c.g.s. units). (Inset : charge distributions.)

Fig. 2



Differential cross sections for scattering at 125 MeV from: (i) Homogeneous distribution with Gaussian cut off

$$\rho = K\frac{1}{1+\exp[(r^2-\alpha^2)/\beta^2]}$$

with $\beta^2=\alpha^2/4$, $\alpha=3{\cdot}91$, $K=0{\cdot}00216$.

(ii) Wine bottle distribution

$$\rho = K\frac{1+r^2/\alpha^2}{1+\exp[(r^2-\alpha^2)/\beta^2]}$$

with $\beta^2=\alpha^2/4$, $\alpha=3{\cdot}68$, $K=0{\cdot}00146(5)$.

(iii) Exponential distribution

$$\rho = K\exp(-r/\alpha)$$

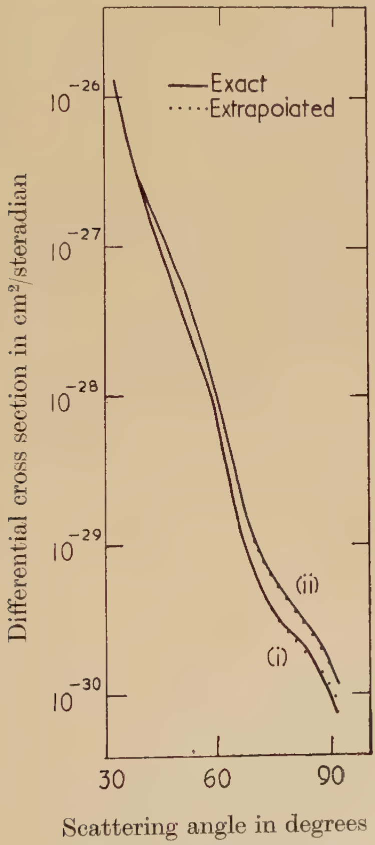
with $\alpha=0{\cdot}99$, $K=0{\cdot}0236$.

(Lengths are measured in units of $\hbar c/E$, energies in units of E.) All distributions have nuclear parameter $r_0=1{\cdot}2$ (in c.g.s. units). (Inset: charge distributions.)

errors are large in the region of disagreement and by drawing the 126 MeV curve in a different way, but still within the range of experimental error, the discrepancies can be decreased.

The fact that experimental curves do not extrapolate very well with the above rule may indicate the presence of small effects not considered here. Order of magnitude estimates of the effects usually considered (Schiff 1953) seem to indicate that they are much smaller than the differences between actual and extrapolated experimental curves. If we assume from our three specific calculations that all reasonable theoretical curves extrapolate with the above rule, and that there are no important effects other than those considered here, we must ascribe the differences mentioned above to experimental errors, and we would then obtain a more reliable fit by trying to fit all energies simultaneously. Even if our assumptions are incorrect and the discrepancies arise from theoretical inadequacies, we

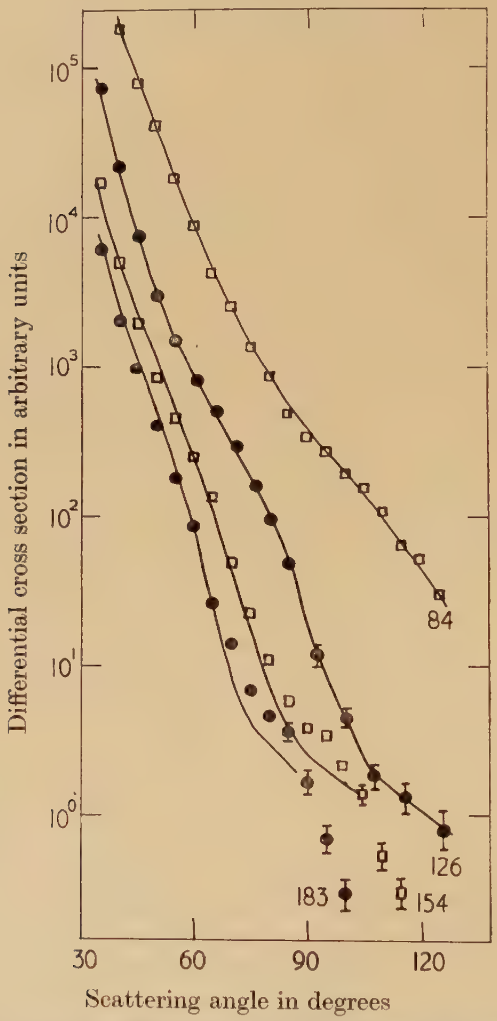
Fig. 3



Theoretical differential cross sections calculated at 180 MeV for the models listed as (i) and (ii) under fig. 2 and curves extrapolated from those of fig. 2 by the rule, eqns. (8) and (9).

have no reason to weight results at a particular energy more heavily than those at another. We illustrate, therefore, in figs. 5–7 comparison of theoretical and experimental results at all energies.

Fig. 4

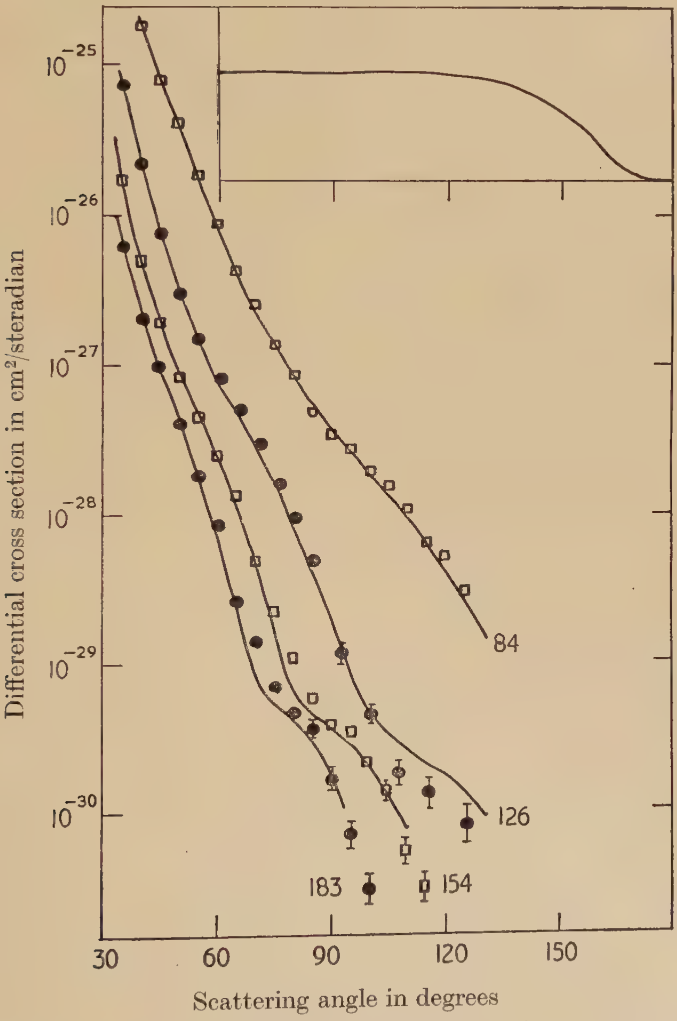


Experimental points at 84, 126, 154 and 183 Mev and the curves obtained by extrapolating to 84, 154 and 183 Mev the line drawn through the 126 Mev points.

We show in figs. 5 and 6 curves of type (*a*) with $\beta^2=\alpha^2/8$ and $\beta^2=\alpha^2/4$ respectively (transition regions of $1{\cdot}72\times10^{-13}$ cm and $3{\cdot}20\times10^{-13}$ cm respectively) compared with experimental points for all experimental energies, and in fig. 7 curves for the wine bottle distribution. The theoretical curves at 84 and 154 Mev in all figures and at 183 Mev in fig. 5 have been obtained by the extrapolation rule.

We see that the general features of the nuclear charge distribution can be ascertained reasonably. The nuclear radius parameter is in the region of $r_0=1{\cdot}2$. For curves of type (*a*), r_0 can be increased by about 0·03, whereas for type (*b*), r_0 can be increased by about 0·06 or decreased by about 0·04 without destroying the fit with experiment. We would say from these results, assuming that the correct model is not radically different from those we have investigated, that r_0 lies in the region 1·16 to 1·26. We see, further, that the actual transition region would seem to lie between the

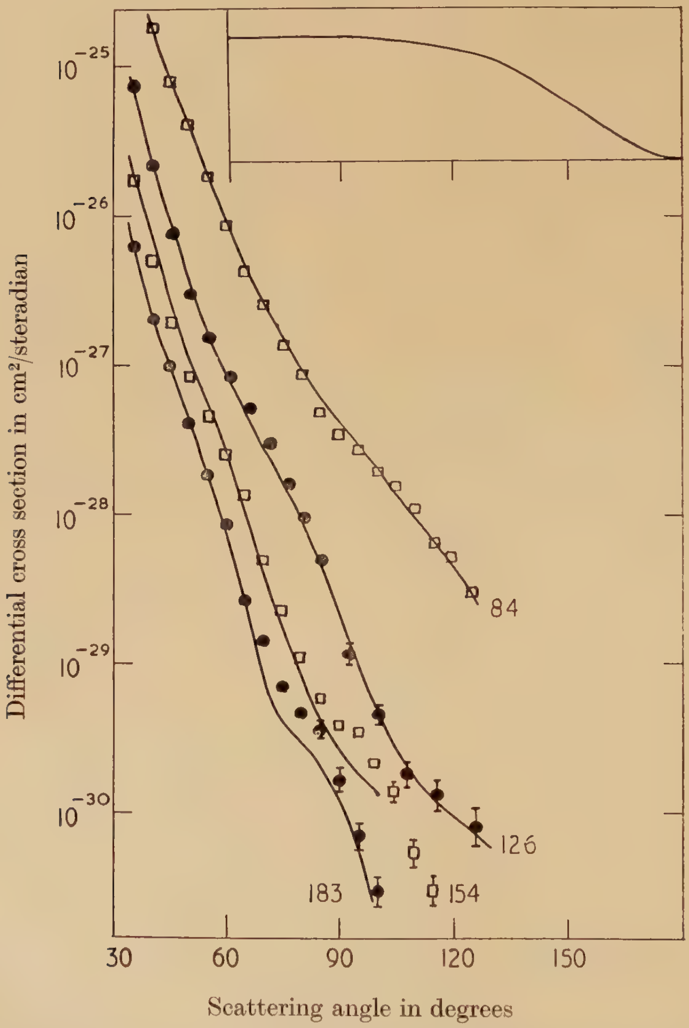
Fig. 5



Theoretical cross sections for the model given below fig. 1 with $\beta^2=\alpha^2/8$, compared with experiment at all energies.

transition regions of the curves, figs. 5 and 6, or between $1{\cdot}72\times10^{-13}$ cm and $3{\cdot}20\times10^{-13}$ cm. Not much can be said about the charge distribution near the origin, since the scattering is relatively insensitive to this region, but a slight wine bottle shape fits the trend of experimental points more easily than a homogeneous distribution in this region.

Fig. 6

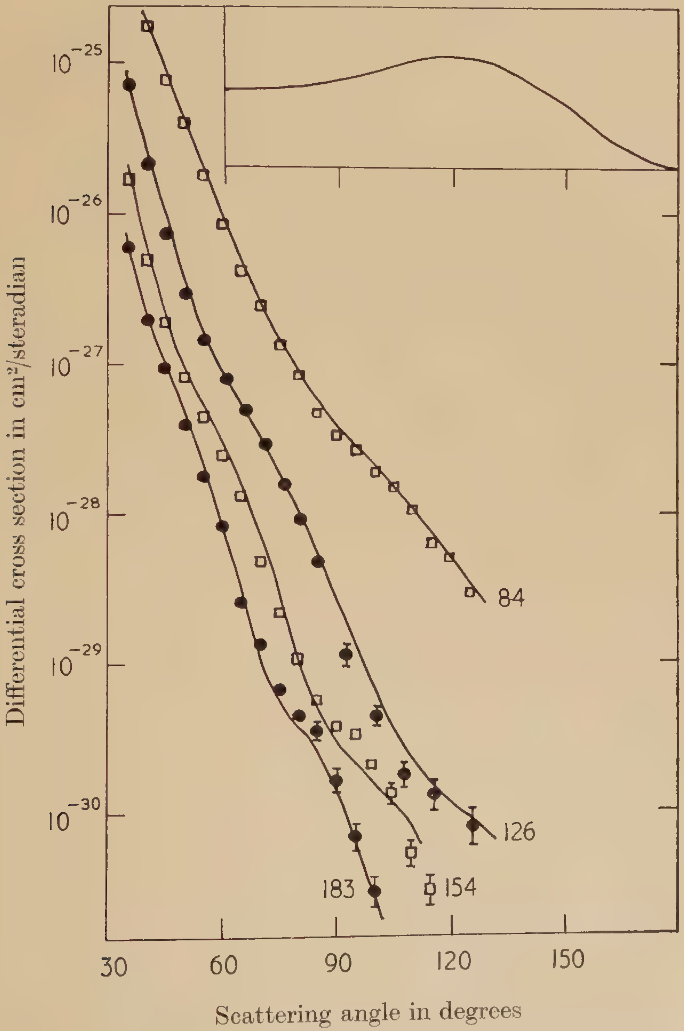


Theoretical cross sections for the model given below fig. 1 with $\beta^2=\alpha^2/4$, compared with experiment at all energies.

We have received results from Dr. Ravenhall and Dr. Yennie which point to similar conclusions in the case of type (*a*) curves, although the details of the charge distributions employed are different ; in particular, an exponential cut off is used instead of a Gaussian cut off, and consequently, numerical values are somewhat different.

We should like to thank the staff of the Cambridge Mathematical Laboratory and in particular the Director Professor M. V. Wilkes for the facilities with which this work was carried out. We should also like to thank Drs. Ravenhall and Yennie for communication of their results before publication and for interesting correspondence, as well as Professor Hofstadter, who sent us the experimental results before publication.

Fig. 7



Theoretical cross sections for the model listed as (ii) below fig. 2, compared with experiment at all energies.

REFERENCES

BRENNER, S., BROWN, G. E., and ELTON, L. R. B., 1954, *Phil. Mag.* (7), **45,** 524.

FESHBACH, H., 1952, *Phys. Rev.*, **88,** 295.

HOFSTADTER, R., HAHN, B., KNUDSEN, A. W., and MCINTYRE, J. A., 1954, *Phys. Rev.*, **95,** 512.

MOTT, N. F., and MASSEY, H. S. W., 1949, *The Theory of Atomic Collisions*, 2nd Ed. (Oxford : University Press).

RAVENHALL, D. G., and YENNIE, D. R., 1954, *Phys. Rev.* **96,** 239.

SCHIFF, L. I., 1953, *Phys. Rev.*, **92,** 988.

YENNIE, D. R., RAVENHALL, D. G., and WILSON, R. N., 1954, *Phys. Rev.*, **95,** 500.

XXII. *The Large-Angle Scattering of μ-Mesons in Iron*

By I. B. McDiarmid
The Physical Laboratories, University of Manchester*

[Received November 22, 1954]

Abstract

The large-angle scattering of μ-mesons in iron has been investigated by a method similar to that reported previously for lead. The measured large-angle distributions are compared with the theories of Molière, Olbert and a modification of Olbert's theory which is based on a single-scattering cross section inferred from recent electron-scattering measurements. At energies from several hundred Mev to about 4 Bev the measured distribution, at large angles, is not consistent with present Coulomb scattering theories in which the effects of the finite nuclear dimensions are included. As in the lead experiment, the number of large angles observed is in rough agreement with Molière's theory for a ' point ' nucleus. Some properties of the so-called anomalous scattering, which follow from the lead and iron experiments, are discussed.

§ 1. Introduction

In a recent paper (McDiarmid 1954, referred to as I), measurements were reported of the μ-meson scattering distribution in a 2 cm lead plate. It was found, in agreement with other workers, that the observed number of large angles for particles with energy in the region 0·6 to 4 Bev was an order of magnitude greater than expected from Olbert's theory of Coulomb scattering for a ' solid ' nucleus. In fact, the large-angle distributions fell very near to those expected from Molière's theory for a ' point ' nucleus.

The purpose of the experiment reported here is to examine (by a method similar to that used in I) the large-angle ' tail ' of the scattering distribution in iron, and from this examination to determine, if possible, the dependence of the large-angle cross section on the scattering nucleus. Again the measurements are compared with the theories of Molière (1948) and Olbert (1952), as well as a modification of Olbert's theory which is based on a more realistic single-scattering cross section for a ' solid ' nucleus. The single scattering law assumed by Olbert is the ' point ' nucleus cross section for angles less than λ/R (λ is the reduced wavelength and R the nuclear radius) and zero for angles greater than λ/R. Although the multiple scattering distribution derived from this is reasonably accurate for large thicknesses and heavy elements, it considerably underestimates the number of large angles in the distribution for small thicknesses and light elements. Hence a simple modification of this theory is given in the Appendix which can be applied to a more or less arbitrary single scattering law.

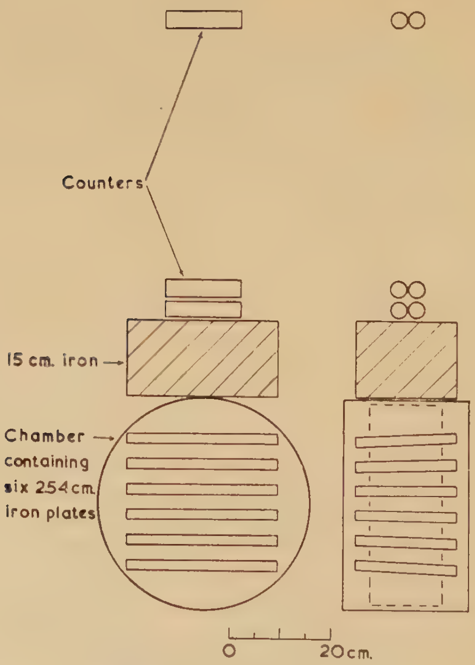
---

* Communicated by H. J. J. Braddick.

§ 2. EXPERIMENTAL ARRANGEMENT

The cloud chamber containing six 2·54 cm iron plates was operated underground at a depth of 26 m water equivalent. The expansions were triggered by the simple counter telescope shown in fig. 1. This counter arrangement, as compared to the one used in the lead experiment, considerably reduced the correction for bias against observing large scattering angles (see I). It also decreased, however, the number of useful pictures taken per hour by slightly more than a factor 3.

Fig. 1



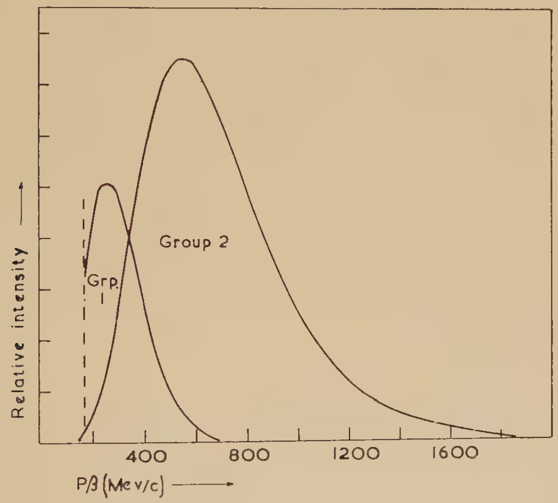
Experimental arrangement.

The rest of the apparatus and the method of measuring the angles were the same as described in I. Again, a check on the accuracy of measurement showed that systematic errors in the angle measurements were less than ±0·2° and that the r.m.s. error in a single measurement of angle was 0·5°.

§ 3. Results

Using the large-angle cut-off mentioned in I the r.m.s. angles of multiple scattering $\langle\theta\rangle_n$ were obtained for about 5000 particles which passed through the chamber and traversed the six iron plates. The frequency distribution of the r.m.s. angles was divided into the following three groups: group 1, 2·6° to 6°; group 2, 1·1° to 2·5°; group 3, $\leqslant$1°. For groups 1 and 2 the $p\beta$ spectra were determined as described in I; these are shown in fig. 2. The large-angle scattering analysis then consists of comparing the number of large angles observed in each group with the number calculated by averaging the different scattering theories given in the Appendix over the $p\beta$ spectra of the corresponding group.

Fig. 2

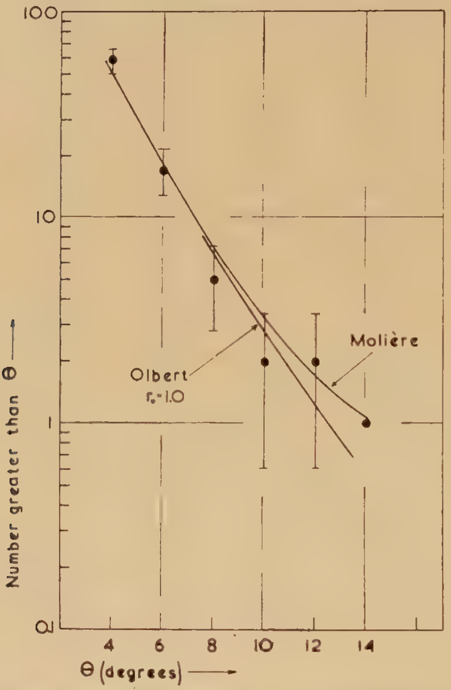


$p\beta$ spectra for groups 1 and 2.

Some of the particles in group 1 lose an appreciable fraction of their energy while traversing the plates and therefore only angles occurring in plates 3 and 4 have been used in the large-angle scattering analysis. The fact that the particles have to traverse the remaining two plates imposes the lower energy limit shown by the dashed line in fig. 2. The particles in group 2 are of sufficiently high energy that ionization loss is not important, hence angles occurring in plates 2 to 5 have been used in the large-angle analysis. In figs. 3 and 4 the measured integral distributions are given for groups 1 and 2 respectively, along with the distributions expected from the theories of Molière and Olbert (for $r_0=1{\cdot}0$, writing $R=r_0\times10^{-13}A^{1/3}$ cm). The present modification of Olbert's theory lies between these two theoretical curves. For the angles and energies considered in both groups the difference between the two theories is small and it is evident that the measurements agree about equally well with either theory. This can be regarded as giving a rough check on the method of analysis used.

The $p\beta$ spectrum (fig. 5) of the particles in group 3 was obtained by subtracting the resultant spectrum of the particles in groups 1 and 2 from the underground spectrum corrected for scattering in the apparatus (see I). The underground spectrum used was derived from the measured sea-level spectrum of Holmes, Owen, Rodgers and Wilson (private communication); although this spectrum is not very different from that assumed in I it is presumably more accurate. The spectrum shown in fig. 5 should be reliable at values of $p\beta$ above about 0·9 Bev/c, but below this there is some uncertainty due to possible errors in the correction applied for scattering in the apparatus and the fact that here the spectrum is derived from the difference of two nearly equal quantities.

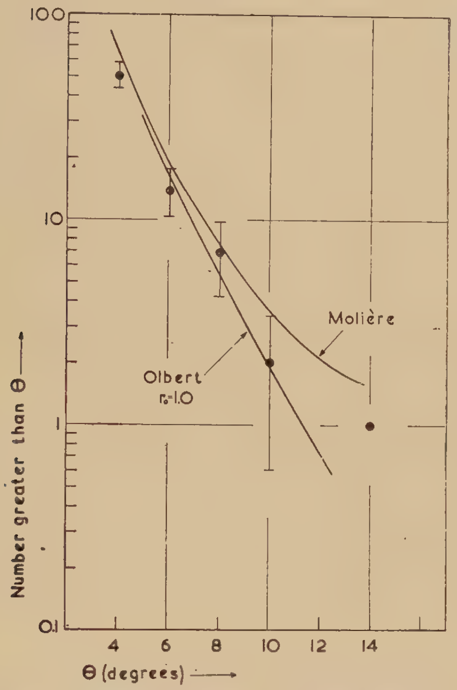
Fig. 3



Measured and calculated integral scattering distributions for group 1. The data represent 226 traversals of a 2·54 cm iron plate.

Again, for particles in this group angles occurring in plates 2 to 5 have been used in the large-angle analysis. In fig. 6 the measured and calculated distributions are given and it can be seen that the measurements are quite inconsistent with either Olbert's theory (with $r_0=1{\cdot}0$) or the present modification of this theory (with $r_0=1{\cdot}1$). As in the lead experiment the measurements are in reasonable agreement with Molière's theory for a 'point' nucleus.

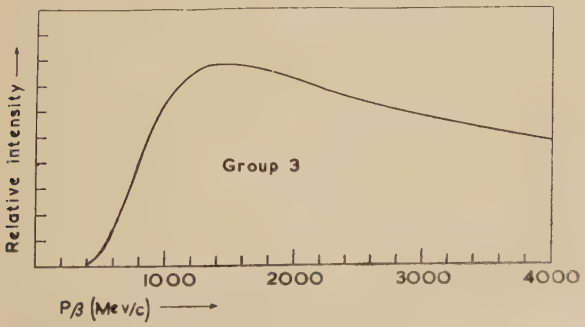
Fig. 4



Measured and calculated integral scattering distributions for group 2. The data represent 1568 traversals of a 2·54 cm iron plate.

An uncertainty exists in the theoretical distributions shown in fig. 6 due to the uncertainty, mentioned above, in the low energy end of the $p\beta$ spectrum of fig. 5 and the fact that low energies contribute appreciably to these distributions. However, this uncertainty is not great enough to affect appreciably the results, e.g. if the number of particles below 0·9 Bev/c is uncertain by a factor 2 (which is certainly an upper limit) then Molière's distribution is uncertain by less than ±20% at all angles and the other distributions by less than ±80% at all angles.

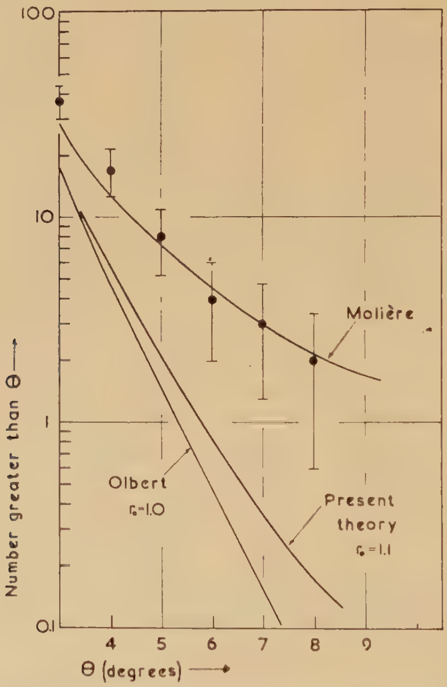
Fig. 5



$p\beta$ spectrum for group 3.

The contribution to the expected distributions due to incoherent (i.e. inelastic) scattering can be estimated from the theory of Gatto (1953). Although this is relatively more important in iron than in lead it is still too small by a factor of nearly 10 to account for the number of large angles observed. In fact, adding this contribution to the ‘ tail ’ of the theoretical distribution (‘ present theory ’) in fig. 6 increases it by about 30% in the region of $\theta=7°$. In connection with incoherent scattering it can be mentioned that no secondary particles were observed to be associated with any of the large-angle events in either lead or iron. If these events are due to incoherent scattering of any kind (Coulomb, short-range forces, etc.) then nuclear disintegration products might be expected in some cases. Whether or not these low energy secondary particles are observed depends on the experimental conditions and it is estimated from the number of large angles observed that, if disintegrations occur, the number observable in both the lead and iron experiments would be considerably less than one. Therefore the fact that none was observed is not evidence against any type of incoherent scattering.

Fig. 6



Measured and calculated integral scattering distributions for group 3. The data represent 18 420 traversals of a 2·54 cm iron plate.

The maximum energy to which the present measurements refer is probably considerably less than might be supposed by looking at the $p\beta$ spectra for groups 3 of this and the lead experiment. Although the mean energy of the particles in both groups is about 8 Bev, as much as half the contribution to the theoretical scattering distributions (Molière's) in figs. 6 and 7 of I, comes from particles with $p\beta$ less than about 2 BeV/c and only 5 to 10% from particles with $p\beta$ greater than about 4 BeV/c. If we assume that the measured scattering has the same energy dependence as Molière's theory (the fact that agreement is obtained with this theory in all groups of both experiments is evidence that it does) then follows that the measurements refer mainly to particles with $p\beta$ less than 4 BeV/c. At energies above this the only information given is an upper limit to the large-angle scattering cross section.

§ 4. CONCLUSIONS AND DISCUSSION

The conclusions drawn from the lead and iron experiments are the following :—

(1) At energies from several hundred MeV to 3 or 4 BeV the μ-meson scattering distribution at large angles in lead and iron is not consistent with present theories of Coulomb scattering (coherent or incoherent) by a ‘ solid ’ nucleus.

(2) At these energies, in both lead and iron, there is an ‘ anomalous ’ large-angle ‘ tail ’ to the distribution, the magnitude of which is in rough agreement with that predicted by Molière's theory for a ‘ point ’ nucleus. In order to account for this anomalous component by the present modification of Olbert's theory the value of the nuclear radius would have to be decreased by a factor 3 (from the value corresponding to $r_0=1\cdot1$).

(3) At energies of a few hundred MeV the present experiments are not capable of distinguishing between the ‘ point ’ and ‘ solid ’ nucleus theories ; at most they set an upper limit to any anomalous scattering component present at these energies.

Although more measurements are required, particularly at low energies, it is still possible from the experiments performed to date to indicate roughly some of the properties of the anomalous scattering. Within the statistical limits of the experiments (in particular the present ones and that of Whittemore and Shutt 1952), all the data are in agreement with the theory of Coulomb scattering for a ‘ point ’ nucleus. This would seem to imply the following : (1) the scattering cross section varies approximately as $1/(p\beta)^2$ or $1/E^2$ (E energy) ; (2) the differential distribution varies as $1/\theta^n$ with $n=3$ to 5 ; (3) the cross section per nucleus varies as Z^2 or A^2 but very probably not as Z or A.

As mentioned in I it does not seem possible at the moment to interpret the measured distributions at large angles. Several authors have suggested that the anomalous scattering indicates a non-Coulomb interaction between fast μ-mesons and nucleons. The difficulty with this suggestion is that any new scattering potential postulated would have to be consistent with the experiments of Conversi *et al.* (1947) and Fitch and Rainwater (1953) on the interaction of stopped μ-mesons with nuclei. Thus, such a potential, as well as predicting the properties of the anomalous scattering given above, would presumably have to approach zero at low energies or else give rise to a very short-range force concentrated at the centre of the nucleus. Neither seems very likely.

On the other hand an explanation in term of Coulomb forces alone must be consistent with the electron scattering data of Hofstadter *et al.* (1953, 1954). Although the momenta, and hence the reduced wavelengths, of the electrons and μ-mesons differ greatly, the maximum momentum transfer observed in both experiments is of the same order ($\sim$ 250 MeV/c). In the Born approximation the scattering cross section depends only on the momentum transfer (the maximum value of this is of course limited by the momentum of the particle), and hence in this approximation no difference would be expected in the two experiments (at least in the coherent scattering). The results of a phase-shift analysis might be expected to depend also on the momentum of the particle (the higher the momentum the more phase shifts have to be included in the analysis). However, such a dependence would not be expected to be great enough to account for the difference between the electron and μ-meson experiments.

Possible the most significant difference between the electron and μ-meson experiments is the fact that only coherent scattering is measured in the former, whereas both coherent and incoherent are measured in the latter. As mentioned above, it is not possible to distinguish between these two types of scattering in the present experiment and this appears to be the case in all μ-meson experiments to date. Thus there is no experimental evidence against an explanation of the anomalous scattering in terms of incoherent Coulomb effects. The existing theory of incoherent Coulomb scattering (Gatto 1953) predicts the observed energy and angle dependence but is inconsistent with the magnitude and possibly the Z dependence of the measured scattering. In this connection a more exact theoretical evaluation of the interaction used by George and Evans (1950), in their analysis of stars produced by μ-mesons, might be useful.

ACKNOWLEDGMENTS

Throughout the course of this work I have benefited from many discussions with Dr. H. J. J. Braddick, Dr. G. N. Fowler, Dr. G. D. Rochester, and Dr. A. W. Wolfendale. The cloud chamber has been operated in conjunction with Mr. B. Leontic, to whom I am indebted.

I am grateful to Professor J. Rainwater for communicating results on the theory of multiple scattering prior to publication.

Finally I am indebted to the Royal Commission for the Exhibition of 1851 for an Overseas Scholarship.

APPENDIX

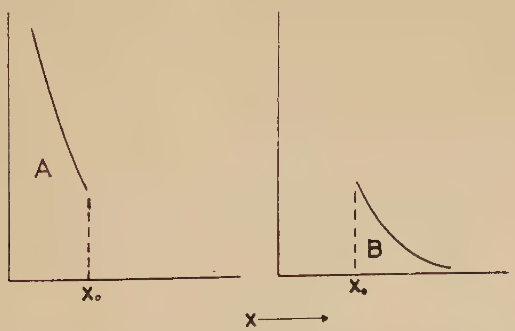
The following is a simple modification of Olbert's theory (1952) which gives the multiple scattering distribution corresponding to a more or less arbitrary single scattering cross section. For reasonable assumptions about the scattering cross section the method gives a distribution which is accurate to about 20% at large angles. The particular single-scattering law used in the present application is one inferred from a Born approximation calculation by Rose (1948), for a uniform nuclear model (U.N.M.). The Born calculation for a U.N.M. gives pronounced maxima and minima in the scattering cross section ; these are not found in either the exact phase shift calculations for 125 MeV electrons (Yennie *et al.* 1953) or in the measured scattering cross section of these particles (Hofstadter *et al.* 1953, 1954). However, a smooth curve lying along the maxima of the cross section given by Rose (for $r_0=1{\cdot}1$ in $R=r_0\times10^{-13}A^{1/3}$) does fit the electron data reasonably well and it is such a single scattering law that is considered here.

In terms of the scattering variable x, defined by Olbert and which refers to projected angles, the probability of a deflection between x and $x+dx$ in consequence of a single collision in a thickness t of scattering material can be written approximately

$$P(x,R)\,dx=\frac{6{\cdot}5}{8}\,\frac{1}{16CG^3}\,\frac{[1-\exp(-4CG^2x^4)]}{x^7}\,dx, \quad . \quad . \quad \text{(A.1)}$$

where the effects of screening, which are not important for the present purposes, have been neglected and where G is defined by Olbert and $C=0{\cdot}5\times10^{-7}(Nt/A)^2Z^4R^4$ for $\beta\sim1$ and R in cm, N=Avogadro's number, t=thickness in g/cm², A=atomic weight, and Z=atomic number. For 2 cm lead $C=1{\cdot}75\times10^{-3}$ ($r_0=1{\cdot}1$) and $G=7{\cdot}15$ for $\beta\sim1$. For 2·54 cm iron $C=3{\cdot}02\times10^{-5}$ ($r_0=1{\cdot}1$) and $G=7{\cdot}30$ for $\beta\sim1$.

Fig. 7



Apart from inherent uncertainties in assuming a cross section given by a smoothed-out Born calculation, eqn. (A.1) is approximate only in that the factor 6·5/8 replaces a complicated function of x obtained when the conversion from total to projected angle is made. This function varies from slightly less than 1 at $x=0$ to $\frac{5}{8}$ at large x so that replacing it with an approximate average value introduces only a slight error.

Following Butler (1950) we now consider a single-scattering law divided at x_0 into two parts A and B as shown in fig. 7, where x_0 is such that the probability of a single collision giving a deflection $x>x_0$ in one traversal of the scattering plate is small. Then to a good approximation the multiple distribution corresponding to this single scattering law can be written $H(x)=H_1(x)+H_2(x)$, where $H_1(x)$ is the multiple distribution for A (i.e. expresses the probability of a resultant deflection x due to all single collisions (x_0) and $H_2(x)$ expresses the probability of a deflection x as the result of *one* single collision greater than x_0 and all single collisions less than x_0. The distribution function $H(x)$ is approximate in that it does not include the effect of more than one single collision greater than x_0. Thus the smaller the probability of such a collision occurring the better will this approximation be.

The multiple distribution due to A, for the case of the single scattering probability for a ‘ point ’ nucleus, is simply the distribution $f(x, x_0)$ given by Olbert. To a good approximation this is also the multiple distribution due to A for the single scattering probability for a U.N.M. given by eqn. (A.1) provided x_0 is chosen small enough so that these two single scattering probabilities do not differ by more than 20 or 30% at $x=x_0$. It is easy to show that such a value of x_0 is given by

$$x_0=\frac{1\cdot 3}{(36CG^2)^{1/4}}. \qquad \text{(A.2)}$$

Therefore, if we choose x_0 as given by eqn. (A.2), then Olbert's distribution will be a good approximation to $H_1(x)$, but at the same time x_0 must be so large that the probability of a deflection $x>x_0$ due to a single collision is small. In general both these conditions are satisfied simultaneously, e.g. for 2 cm lead and 2·54 cm iron $(r_0=1\cdot 1)$ eqn. (A.2) gives $x_0\sim 1$ and ~ 3 respectively and the probability of $x>x_0$ from eqn. (A.1) is less than 0·01 and 0·001 respectively.

The function $H_2(x)$ gives the probability of a resultant deflection $x=x'+x''$ due to one single collision $x'>x_0$, the probability of which is given by eqn. (A.1), and a deflection x'' due to all single collisions less than x_0, the probability of which is given by Olbert's distribution. Thus $H_2(x)$ is simply the joint distribution of these two probability functions, namely,

$$H_2(x)=\int_{-\infty}^{\infty} f(x''=x-x', x_0)P_0(x', R)\,dx',$$

where $f(x''=x-x', x_0)$ is Olbert's function, for the appropriate value of x_0, with the variable replaced by $x-x'$, and $P_0(x', R)$ is given by eqn. (A.1) for $x'>|x_0|$ and is zero for $x'<|x_0|$.

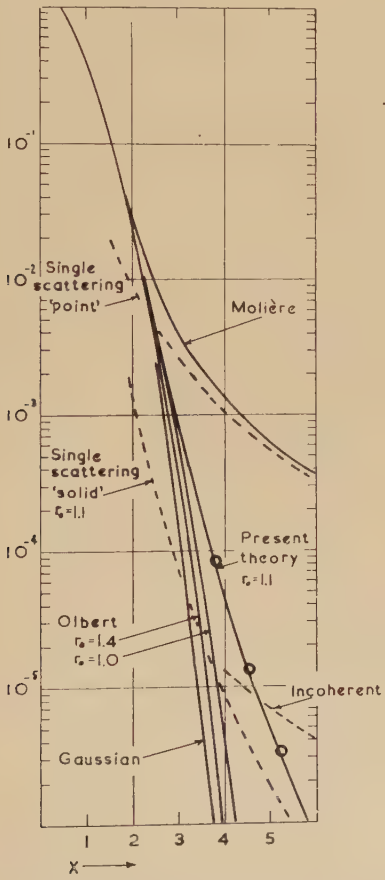
Finally we have

$$H(x)\,dx=f(x, x_0)\,dx+dx\int_{-\infty}^{\infty} f(x''=x-x', x_0)P_0(x', R)\,dx'. \qquad \text{(A.3)}$$

No attempt has been made to evaluate the integral in eqn. (A.3) analytically. Given Olbert's distribution for the appropriate value of x_0, $H(x)$ may then be obtained for a particular case by a numerical computation.

In fig. 8 the distribution function $H(x)$, for 2 cm lead and $r_0=1{\cdot}1$, is compared with Molière's distribution for a 'point' nucleus and Olbert's distribution for $r_0=1{\cdot}4$ and $1{\cdot}0$. These distributions as shown have been multiplied by a factor 2 and thus refer to absolute deflections, i.e. positive and negative. It can be seen that Olbert's distribution, even for the smaller value of r_0, is only a fair approximation to $H(x)$ at large x, and that the distribution obtained simply by adding the single-scattering probability eqn. (A.1) (curve marked 'single-scattering solid') to the small-angle Gaussian approximation would be too low by as much as a factor of 5 in the region $x=3$ to 4.

Fig. 8

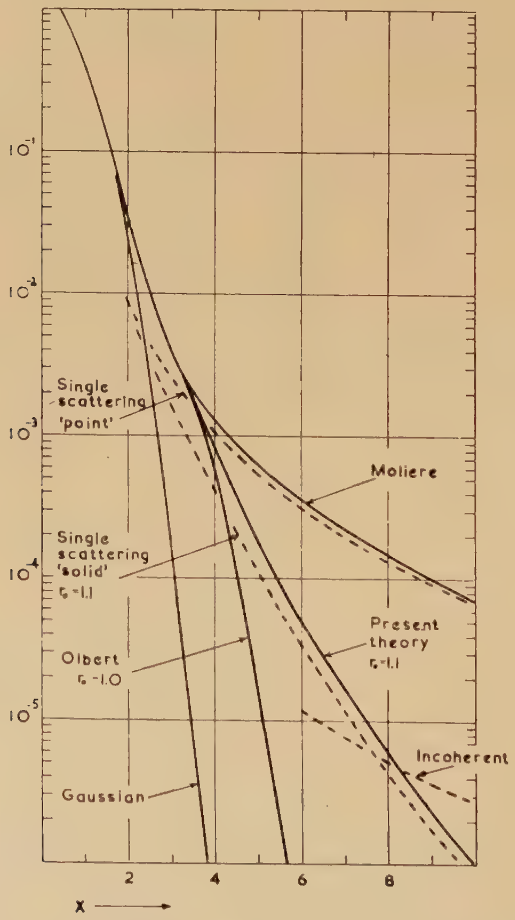


Coulomb scattering distributions for 2 cm of lead. x is the projected angle of scattering ϕ in units of $(2QG)^{1/3}$ (see Olbert 1952). The points shown by circles have been obtained from the multiple scattering theory of Cooper and Rainwater (1954).

In fig. 9 the distribution function $H(x)$ is given for 2·54 cm iron and $r_0=1{\cdot}1$. Comparing this with the other distributions shown it is evident that Olbert's greatly underestimates the number of large deflections (i.e. by a factor of 10 at $x=5$) but, as in the case of the 'point' nucleus distributions, a fair approximation to $H(x)$ would be simply the Gaussian approximation at small x plus the single-scattering probability at large x. In general, the smaller the slope of the single-scattering curve the better is this approximate distribution given by the sum of the Gaussian and single-scattering tail.

Also shown in figs. 8 and 9 are the relative contributions due to incoherent scattering for large energy transfers estimated for the theory of Gatto (1953).

Fig. 9



Coulomb scattering distributions for 2·54 cm of iron.

Recently Cooper and Rainwater (1954 and private communication) using an extension of Molière's theory, have obtained a multiple scattering distribution which can also be applied to an arbitrary single scattering cross section. These authors give a distribution for 2 cm of lead for a single scattering law very similar to the one used here (for $r_0=1{\cdot}1$) and hence a comparison is possible. The scattering variable used in ϕ/ϕ_0 which, for the case of 2 cm lead and $r_0=1{\cdot}1$, is related to Olbert's variable x by $x=0{\cdot}75\phi/\phi_0$. To convert from the Cooper–Rainwater distribution to the one given in fig. 8 it is necessary to multiply their distribution by a factor 1/0·75 in order to renormalize and also by a factor 2 because the distributions in fig. 8 are for absolute angles. Because of slight differences between the single scattering law used by Cooper and Rainwater and eqn. (A.1), their distributions (single and multiple) at large angles have to be multiplied by a further factor $9/12\times6{\cdot}5/8$ before being compared with those of fig. 8. The factor 9/12 arises from a difference in the assumed form factors (see Schiff 1953) and the factor 6·5/8 from the fact that Cooper and Rainwater tke the same single scattering law for projected as for total angle. Therefore, converting from ϕ/ϕ_0 to x and multiplying their distribution (at large ϕ/ϕ_0) by the above factors gives the points shown in fig. 8. The agreement is seen to be excellent.

REFERENCES

BUTLER, S. T., 1950, *Proc. Phys. Soc.* A, **63,** 599.
CONVERSI, M. E., PANCINI, E., and PICCIONI, O., 1947, *Phys. Rev.*, **71,** 209.
COOPER, L. N., and RAINWATER, J., 1954, *Phys. Rev.*, **95,** 1107.
FITCH, V. L., and RAINWATER, J., 1953, *Phys. Rev.*, **92,** 789.
GATTO, R., 1953, *Nuovo Cimento*, **10,** 1559.
GEORGE, E. P., and EVANS, J., 1950, *Proc. Phys. Soc.* A, **63,** 1248.
HOFSTADTER, R., FECHTER, H. R., and MCINTYRE, J. A., 1951, *Phys. Rev.*, **92,** 978; 1954, *Ibid.*, **95,** 512.
MCDIARMID, I. B., 1954, *Phil. Mag.*, **45,** 933.
MOLIÈRE, G., 1948, *Zeits. f. Naturforschung*, 3a, 78.
OLBERT, S., 1952, *Phys. Rev.*, **87,** 319.
ROSE, M. E., 1948, *Phys. Rev.*, **73,** 279.
SCHIFF, L. I., 1943, *Phys. Rev.*, **92,** 988.
WHITTEMORE, W. L., and SHUTT, R. P., 1952, *Phys. Rev.*, **88,** 1312.
YENNIE, D. R., WILSON, R. N., and RAVENHALL, D. G., *Phys. Rev.*, **92,** 1325.

XXIII. *Collision Processes in Meteor Trails*

By H. S. W. MASSEY and D. W. SIDA
University College, London*

[Received November 22, 1954]

SUMMARY

Extensive data on the probability that an evaporated meteor atom will produce an electron in an ionizing collision before being brought to rest are now available. To obtain from these data values of the ionization cross section for impacts between meteor atoms and atmospheric molecules the momentum loss cross section for elastic impacts of this kind must be known. These latter cross sections are calculated for roughly representative collisions in which the interaction between a meteor atom and a gas molecule is taken to be that between calcium and neon atoms calculated statistically. Values for the ionization cross sections are then derived and found to be consistent with the recent theoretical considerations of Bates and Massey (1954).

The diffusion of ionization in a meteor trail is also discussed in relation to known observed data on the mobility of positive ions in gases.

§ 1. INTRODUCTION

THE study of the optical and ionization phenomena associated with the development of a meteor trail has progressed so far that it is possible to provide a quite detailed interpretation of the events involved in terms of certain atomic collision coefficients. The values of these coefficients averaged over the various meteor constituents may be obtained from an analysis of the meteor data and the question arises as to whether they are consistent with atomic collision theory and with direct measurement.

The ionization produced by a meteor is due to the impact of atoms evaporated from the meteor with atmospheric atoms and molecules. From the meteor data it is possible to derive the chance that a meteor atom will produce an electron before it is brought to thermal equilibrium with the surrounding atmosphere. For comparison with collision theory and with direct experimental data it is necessary to relate this chance to the appropriate collision cross sections. The chance can be expressed roughly as Q_i/Q_d where Q_i is the mean cross section for ionization and Q_d for momentum loss (or diffusion) in a collision between a meteor atom and an atmospheric molecule. If Q_d can be calculated Q_i may be obtained and compared with expectation. Similar considerations apply to the emission of light by a meteor, the chance that a meteor atom will emit a quantum with frequency in the observed range being given by Q_{ex}/Q_d where Q_{ex} is

---

* Communicated by the Authors.

the mean cross section for an inelastic collision which leads ultimately to emission of an appropriate quantum.

In this note we report the results of a calculation of Q_d for a roughly representative collision, the impact of calcium atoms with neon atoms and discuss briefly the derived values of Q_i.

For the above application the cross section Q_d is required for atoms with energies of 100 ev or more but the rate of decrease of ionization density in a meteor trail is determined largely by the rate of diffusion for positive ions with thermal velocities. These ions are mainly ionized meteor atoms and their diffusion coefficients may be estimated quite closely from observed data on the mobility of alkali metal ions in nitrogen. The comparison of these estimated values with those derived from meteor observations is discussed.

§ 2. PROBABILITY OF IONIZATION OR EXCITATION BY A METEOR ATOM IN TERMS OF COLLISION CROSS SECTIONS

Let $I(E, \theta)d\omega$ be the differential cross section for an elastic collision between a meteor atom of kinetic energy E and a gas molecule, in which the direction of relative motion is turned through an angle θ. If the masses of the colliding systems are nearly equal the kinetic energy lost by the meteor atom as a result of the collision is given closely by

$$\tfrac{1}{2}E(1-\cos\theta).$$

Hence in proceeding a distance δs in an atmosphere containing n gas molecules per c.c. the amount of energy lost in elastic collisions is given by

$$\tfrac{1}{2}nE\delta s\int(1-\cos\theta)I(E,\theta)d\omega=\tfrac{1}{2}nEQ_d(E)\delta s \quad . \quad . \quad . \quad . \qquad (1)$$

where

$$Q_d(E)=\int(1-\cos\theta)I(E,\theta)\,d\omega, \quad . \quad . \quad . \quad . \qquad (2)$$

is usually known as the momentum loss cross section. We now have

$$\frac{dE}{ds}=-\tfrac{1}{2}nEQ_d(E)-n\overline{E}_iQ_i-n\overline{E}_{ex}Q_{ex} \quad . \quad . \quad . \quad . \qquad (3)$$

where Q_i, Q_{ex} are cross sections for ionization and excitation collisions between meteor atoms of energy E and gas molecules and $\overline{E}_i$, $\overline{E}_{ex}$ the mean energies lost in these respective collisions. If δN_e is the number of electrons produced in ionizing collisions in passing the distance δs we have

$$\frac{dN_e}{ds}=nQ_i.$$

Using (3) we have

$$\frac{dN_e}{dE}=-Q_i/(\tfrac{1}{2}EQ_d+\overline{E}_iQ_i+\overline{E}_{ex}Q_{ex}). \quad . \quad . \quad . \quad . \quad . \qquad (4)$$

This gives for β, the chance that the atom will produce an electron before its energy is reduced below the ionization threshold $(2E_i)$,

$$\beta=\int_{2E_i}^{E}Q_i(E)dE/(\tfrac{1}{2}EQ_d+\overline{E}_iQ_i+\overline{E}_{ex}Q_{ex}). \quad . \quad . \quad . \quad . \qquad (5)$$

In practice the denominator can be put equal to $\frac{1}{2}EQ_d$ without serious error. This is because Q_d is at least comparable with Q_i and Q_{ex} when $E \gg \overline{E}_i$ or $\overline{E}_{ex}$ and is much greater when $E \simeq E_i$. Actually Q_i will in general be small for values of E considerably greater than $2E_i$. If E_t is the value of E below which Q_i can be neglected we have approximately

$$\beta=2\int_{E_t}^{E}\frac{Q_i(E)dE}{Q_d(E)E} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

A similar expression holds for the chance of excitation.

In practice, of course, there are many kinds of meteor atoms involved and at least two kinds of atmospheric molecules. We must then regard the cross sections Q_i, Q_d which appear in (6) as suitably weighted means over all the collisions of a particular kind which can occur. It is also assumed that β is less than unity so that consideration does not have to be given to the changes in the cross sections which result after ionization of the meteor atom occurs. The interpretation of ionization effects is essentially simpler than that of excitation for the ionization due to a meteor is directly observed whereas excitation can be only studied indirectly in terms of light emission. Many complicating factors are thereby introduced.

To carry out a definite calculation we take calcium as a representative meteor atom. It is not possible to evaluate Q_d for collisions of calcium atoms with either N_2 or O_2 molecules. The detailed calculations were therefore carried out for collisions with an atom ; that of neon. This has the advantage of possessing a closed shell of electrons for which the method of calculation is most suitable and an atomic number close to that of nitrogen and oxygen. It also has the advantage that experimental verification of the calculated cross sections is at least a possibility. In the formula (1) it was assumed that the colliding systems had approximately the same mass. This feature will still be retained by taking as a representative collision one in which the interaction is that between a calcium and a neon atom but with the masses equal to that of calcium.

§ 3. CALCULATION OF MOMENTUM LOSS CROSS SECTION

It is well known that for encounters between systems of atomic mass classical mechanics is adequate except for very small angles of scattering. Owing to the weighting factor $(1-\cos\theta)$ which appears in the formula (2) for Q_d these angles make a negligible contribution. We therefore have (Massey and Burhop 1952), when the masses of the two systems are equal,

$$I(E,\theta)=\frac{p}{\sin\theta}\frac{dp}{d\theta}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where

$$\theta=\pi-2\int_{r_0}^{\infty}\frac{dr}{r\phi(r)},$$

$$\phi(r)=\left\{\frac{r^2}{p^2}-1-\frac{2r^2}{p^2E}V(r)\right\}^{1/2}.$$

$V(r)$ is the interaction energy and r_0 is the largest positive zero of $\phi(r)$.

The interaction energy between two atoms may be calculated to a good approximation by a statistical method due to Lenz (1932), Jensen (1932) and Gombàs (1935). This treatment assumes that the electrostatic potential and electron charge density of the two systems are very little distorted by the interaction. The total potential and charge density are then obtained by superposition. The total potential energy is then given by the usual electrostatic formula and the total kinetic energy, which must include allowance for the increase due to the Pauli Principle, by the statistical formula for a Fermi gas

$$E_{\text{kin}}=\frac{4\pi}{5}\left(\frac{3}{8\pi}\right)^{5/3}\frac{h^2}{m}\rho^{5/3}$$

and a term correcting for exchange relations between electrons with parallel spins

$$E_{ex}=-\frac{3}{2}\left(\frac{3}{8\pi}\right)^{1/3}e^2\rho^{4/3}.$$

m is the electron mass and ρ the electron charge density.

In this way the interaction energy is obtained in the form

$$\begin{aligned}V(r)=&\frac{Z_1Z_2e^2}{r}-e\int\left\{\frac{Z_1\rho(r_2)}{r_1}+\frac{Z_2\rho(r_1)}{r_2}\right\}d\tau-\frac{1}{2}\int\{V_e(r_1)\rho(r_2)+V_e(r_2)\rho(r_1)\}d\tau\\&+\frac{(4\pi)^{1/3}}{5e^{5/3}}\left(\frac{3}{8\pi}\right)\frac{h^2}{m}\int[\{\rho(r_1)+\rho(r_2)\}^{5/3}-\{\rho(r_1)\}^{5/3}-\{\rho(r_2)\}^{5/3}]d\tau\\&-\frac{3}{2}\left(\frac{3}{8\pi}\right)^{1/3}\frac{e^{2/3}}{(4\pi)^{1/3}}\int[\{\rho(r_1)+\rho(r_2)\}^{4/3}-\{\rho(r_1)\}^{4/3}-\{\rho(r_2)\}^{4/3}]d\tau.\end{aligned} \qquad (8)$$

r_1, r_2 are measured respectively from the nuclei of the two atoms, of respective charges Z_1e, Z_2e so that $\mathbf{r}_1-\mathbf{r}_2=\mathbf{r}$. $V_e(r_1)$, $V_e(r_2)$ are the respective electrostatic potentials due to the electron clouds in the two atoms, $\rho(r_1)$, $\rho(r_2)$ the respective electron charge densities.

This statistical formula is probably most satisfactory for use in discussing collisions between systems with closed shells. This is not quite true in our case but the part played by the two outer electrons in the calcium atom will be very small when the kinetic energy of relative motion is as high as 100 ev as in the meteor encounters, and the effective interaction is that between Ca^{++} and Ne for which (8) should be a good approximation.

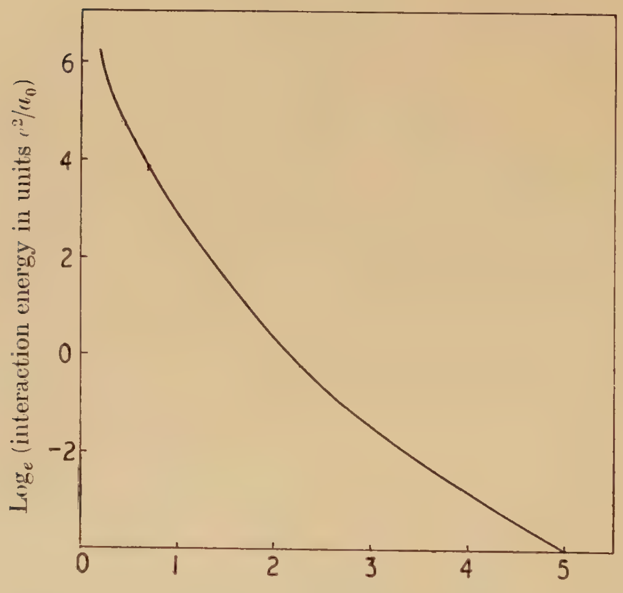
Self-consistent fields for Ca (Hartree and Hartree 1938) and Ne (Brown 1933) are available from which ρ and V_e in (8) may be obtained. Using these data $V(r)$ was calculated from (8) using numerical integration. It is illustrated in fig. 1. Further numerical integration could then be carried out to obtain $I(E, \theta)$ from (7). Calculations were carried out for

$E=50$, 100, 500, 1000 ev and the results are illustrated in fig. 2. Finally Q_d was obtained for these three energies, the values being given in table 1.

Table 1. Momentum Loss Cross Section for Collisions between Ca and Ne Atoms

E (actual) ev	E (assuming masses equal) ev	Q_d (in units πa_0^2)
50	75	4·39
100	150	3·20
500	750	1·34
1000	1500	0·77

Fig. 1



Interaction energy between calcium and neon atoms, calculated by statistical methods.

§ 4. Ionization Probability for Meteor Atoms

If we assume initially that the ionization cross section for impacts between a meteor atom of kinetic energy E and a gas molecule is effectively independent of E for $E > E_t$, the threshold energy in (6), we have that

$$\beta=2Q_i\int_E^{E_t}\frac{dE}{EQ_d(E)}.$$

Over the energy range of interest Q_d can be represented with sufficient accuracy by the form

$$Q_d=Q_0(\log E/C)^2,$$

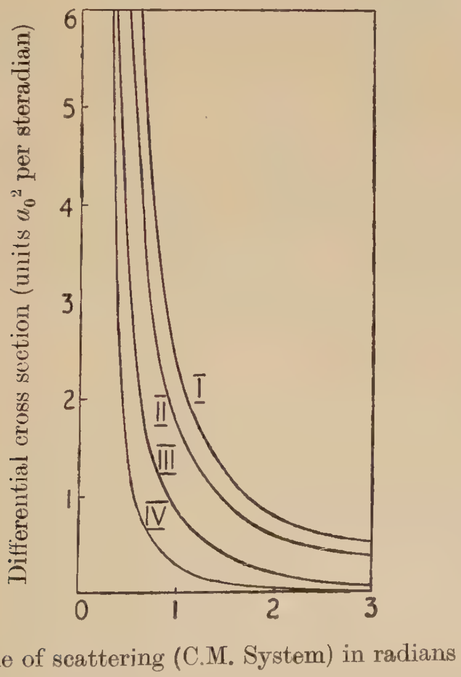
where $Q_0=0{\cdot}165\ \pi a_0^2$ and $C=12\,800$ ev. We find the following values of Q_i/β for different initial energies and two reasonable values for the threshold energy E_t.

Table 2

Initial energy of meteor atom ev	Q_i/β in units πa_0^2 $E_t=20$ ev	 $E_t=10$ ev
1000	0·35	0·33
500	0·54	0·49
100	1·61	1·25

It will be seen that, to be consistent with our assumption that Q_i is effectively constant over the energy range from E_t to 1000 ev, the probability of ionization β should increase with the velocity v of the meteor atoms as v^n where $n=1{\cdot}3$ or $1{\cdot}2$ according as E_t is taken as 20 or 10 ev. The most recent observed data (Kaiser 1954) give $n=0{\cdot}5\pm0{\cdot}5$.

Fig. 2



Differential cross sections for elastic collisions between calcium and neon atoms. Curves I, II, III, IV refer to collisions in which the neon atoms are at rest and the calcium atoms have kinetic energies of 50, 100, 500 and 1000 ev respectively.

This suggests that it is likely that Q_i increases somewhat as v decreases. According to the theoretical considerations of Bates and Massey (1954) Q_i varies approximately as v^{-1} over this energy range so that the calculation of β was repeated with this assumption. It was then found that $n=0{\cdot}87$

($E_t = 20$ ev) or 0·57 ($E_t = 10$ ev), either of which is certainly not in conflict with the present observed data.

Assuming that the theoretical variation of Q_i with v is correct the value of Q_i comes out to be $Q_i^0(1000/E)^{1.2}\,\beta_0\pi a_0^2$ where $Q_i^0 = 0{\cdot}35$ or $0{\cdot}33$ according as $E_t = 20$ or 10 ev respectively. E is the kinetic energy of the meteor atom in ev and β_0 is the value of the ionization probability for meteor atoms of 1000 ev energy. β_0 is not known very accurately but a value of 0·2 seems to be a reasonable estimate. This gives absolute values of Q_i^0 which fall within the range of values expected theoretically. Thus, according to Bates and Massey (1954), the ionization cross section due to impact of atoms A and B is given approximately by

$$Q_i = (3{\cdot}8 \times 10^{-15}\, gM^{1/2}R_x^{\,3}/E^{1/2}t)\pi a_0^{\,2}$$

where g is a statistical weight factor of order unity, M is the reduced mass of the colliding systems (with the proton mass as unit), R_x is the separation at which the potential energy curve along which A and B initially approach intersects a suitable one for interaction of A and B^+, E is the energy of relative motion in ev and t is the lifetime in seconds for auto-ionization of the complex AB when the separation is less than R_x. Values of t between 10^{-13} and 10^{-15} sec are reasonable. Assuming these and taking g as unity we find that the meteor data require that R_x should lie in the range $0{\cdot}3a_0$ to $2{\cdot}0a_0$, which is not unreasonable.

We may therefore summarize the situation as one in which the theory is consistent with the observations but there is a great deal of flexibility in the comparison which can only be removed by laboratory experiments on the ionization of atoms by atom impact. They are very difficult to carry out but would seem to be worthwhile.

§ 5. AMBIPOLAR DIFFUSION COEFFICIENT IN A METEOR TRAIL

The rate of decay of the radio echo from a meteor trial gives a measure of the rate of decrease of electron concentration in the trail. This decay is due to ambipolar diffusion of ions and electrons, to recombination and to electron attachment. Analysis of the evidence suggests that the latter modes of decay are usually unimportant (Kaiser 1953), the main loss being by ambipolar diffusion. From a study of the decay rates of echo from the Arietid meteors Kaiser finds that the most probable value of the ambipolar coefficient D at the height concerned (92·5 km), where the particle density n is 7×10^{13}/c.c., is $2{\cdot}3 \times 10^4$ c.g.s. units.

An ionized meteor atom of mass M will come into thermal equilibrium with atmospheric molecules in a time t of order

$$\frac{(2M)^{1/2}}{n}\int_{E_i}^{E_t}\frac{dE}{E^{3/2}Q_d(E)}\,.$$

where E_i is the initial kinetic energy of the ion and E_t the mean thermal energy of a gas molecule. Substitution of calculated values of Q_d shows that this time is between 10^{-4} and 10^{-3} sec at 92·5 km altitude. For electrons the corresponding time is of order

$$\frac{M}{n}\left(\frac{1}{8m}\right)^{1/2}\int_{E_e}^{E_t}\frac{dE}{E^{3/2}Q_d^{\,e}(E)},$$

where $Q_d^{\,e}$ is the momentum loss cross section for encounters of electrons with gas molecules. This may be taken as about 10^{-15} cm² and is very roughly independent of energy. If $E_e \gg E_t$ we find that a time of about 0·05 sec is necessary for the electrons to come to thermal equilibrium at 92·5 km altitude. This is not far below the echo decay times observed for the Arietid meteors but it is probably quite a good approximation to assume that both the electrons and ions have come to thermal equilibrium with the surrounding atmosphere. Under these conditions and, provided there is no magnetic field present, the ambipolar diffusion coefficient $D=2D_i$, where D_i is the diffusion coefficient of the positive ions. When such a field is present the situation is more complicated. If fluctuations in ion and electron concentration are ignored the diffusion of electrons transverse to the magnetic field of strength H is reduced by the factor $\nu_e^2/(\nu_e^2+\omega^2)$ where ν_e is the electron collision frequency and $\omega=eH/mc$. It has been pointed out by Kaiser (1953) that the effect of this factor would be to reduce the electron diffusion coefficient D_e transverse to the field to a value comparable with D_i at a height of about 100 km. When $D_i=D_e$ the ambipolar diffusion coefficient D becomes equal to D_i instead of $2D_i$. We therefore expect that for the Arietid meteors the mean ambipolar coefficient D should be between D_i and $2D_i$.

D_i can be written in terms of the ionic mobility k_i by the relation

$$D_i=k_i\left(\frac{\kappa T}{e}\right),$$

where κ is Boltzmann's constant. The mobilities of alkali ions in nitrogen of such purity that no clustering takes place have been measured by Powell and Brata (1932). It is unlikely that the mobility in oxygen will be very different. Using the observed values we may determine D_i for the conditions prevailing for the Arietid echoes. Allowance may be made for the low temperature (205°K) at 92·5 km as the ion diffusion coefficient may be assumed to be proportional to the absolute temperature for constant gas density. We find for Na^+, K^+, Rb^+ and Cs^+ ions respectively that $D_i=2{\cdot}37$, $2{\cdot}10$, $1{\cdot}86$ and $1{\cdot}83\times10^4$ c.g.s. units respectively. In view of the fact that for singly charged ions in the same gas the mobility varies approximately as $(1+M_1/M_2)^{1/2}$, where M_1 is the mass of a gas molecule and M_2 of an ion, it is to be expected that this range of values of D_i includes all those likely to arise in a meteor trail. The agreement with the mean value derived from the Arietid meteor echo decay, which is between $1{\cdot}15$ and $2{\cdot}3\times10^4$ c.g.s., can therefore be regarded as reasonably satisfactory.

ACKNOWLEDGMENTS

We are grateful to Dr. T. R. Kaiser for many stimulating discussions on meteoric phenomena. One of us (D. W. S.) also wishes to acknowledge the award of a Maintenance Grant from the Department of Scientific and Industrial Research.

REFERENCES

BATES, D. R., and MASSEY, H. S. W., 1954, *Phil. Mag.*, **45,** 111.
BROWN, F. R., 1933, *Phys. Rev.*, **44,** 214.
GOMBÀS, P., 1935, *Zeits. f. Phys.*, **94,** 473.
HARTREE, D. R., and HARTREE, W., 1938, *Proc. Roy. Soc.* A, **164,** 167.
JENSEN, H., 1932, *Zeits. f. Phys.*, **77,** 722.
KAISER, T. R., 1953, *Advances in Physics*, **2,** 495 ; 1954, Private Communication.
LENZ, W., 1932, *Zeits. f. Phys.*, **77,** 713.
MASSEY, H. S. W., and BURHOP, E. H. S., 1952, *Electronic and Ionic Impact Phenomena* (Oxford : Clarendon Press), Ch. VII.
POWELL, C. F., and BRATA, L., 1932, *Proc. Roy. Soc.* A, **138,** 117.

XXIV. *Oscillatory Motion of a Viscous Liquid in a Thin-Walled Elastic Tube—I: The Linear Approximation for Long Waves*

By J. R. WOMERSLEY
Physiology Dept., St. Bartholomew's Medical College*

[Revised MS. received October 19, 1954]

§ 1. INTRODUCTION AND SUMMARY

THE problem of determining the motion of a liquid in an elastic tube when subjected to a pressure-gradient which is a periodic function of the time arises in connection with the flow of blood in the larger arteries (Helps and McDonald 1954, Womersley 1954). Attempts have been made in the past to measure the rate of flow in the aorta and femoral artery of the dog and rabbit (Shipley, Gregg and Schroeder 1943) and to relate these observations to the varying pressure. In the absence of an adequate mathematical theory, these were not very successful. More recent determinations by direct observation of the motion through the translucent arterial wall, using high-speed cinematography (Helps and McDonald 1954, Womersley 1954) have been accompanied by measurements, not only of pulse-pressure, but of pressure-gradient. Fair agreement has been shown between the observed rates of flow and a simple solution for oscillatory motion of a viscous liquid in a tube with rigid walls (Womersley, in press).

In this paper the corresponding solution for a thin-walled elastic tube is given, it being assumed that the effect of the inertia terms in the equations of viscous fluid motion can be neglected. An approximate correction for the effect of the inertia terms will be presented in Part II of this communication.

It is shown that, when the liquid contained in the tube is viscous, the pressure-wave cannot be propagated without distortion. Not only is the motion damped, but the wave-velocity rises as the frequency increases. For constant frequency, the wave-velocity rises as the viscosity of the liquid decreases, tending to an asymptotic value equal to that for a perfect fluid given by Lamb (1898). It is also shown that the longitudinal oscillation of the walls of the tube (caused by the viscous drag on its inner surface) is important in determining the rate of flow, which may be 10% greater than that in a rigid tube under the same pressure-gradient.

---

* Communicated by Sir Geoffrey Taylor

§ 2. THE SOLUTION FOR THE RIGID TUBE

The simple solution for the oscillatory motion of a viscous liquid in a rigid tube, under a simple-harmonic pressure gradient, was given by Lambossy (1952) who gave the formulae for velocity and viscous drag. He did not apply his result to the motion of the blood in arteries, being solely concerned with the effect of the viscous drag on the frequency response of pressure recording instruments. The author obtained the same result independently, in a different form, and derived the expression for the rate of flow. This result has been used to predict rates of flow from observed pressure-gradient (Womersley, in press). This solution is repeated here for completeness. Let R be the radius of the tube, w the velocity along the tube, Ae^{int} the pressure-gradient, μ the viscosity of the liquid, ρ the density of the liquid, $\nu=\mu/\rho$ the kinematic viscosity.

The equation of motion is

$$\frac{\partial^2 w}{\partial r^2}+\frac{1}{r}\frac{\partial w}{\partial r}-\frac{1}{\nu}\frac{\partial w}{\partial t}=\frac{A}{\rho}e^{int}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Let $y=r/R$ and $w=ue^{int}$ and let the non-dimensional quantity $R\sqrt{(n/\nu)}$ be denoted by α. The equation for u is

$$\frac{d^2u}{dy^2}+\frac{1}{y}\frac{du}{dy}-i\alpha^2u=\frac{AR^2}{\mu} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

and therefore
$$w=\frac{AR^2}{\mu}\cdot\frac{1}{i^3\alpha^2}\left\{1-\frac{J_0(\alpha i^{3/2}y)}{J_0(\alpha i^{3/2})}\right\}e^{int}. \quad . \quad . \quad . \quad . \quad (3)$$

In Lambossy's paper he separates this into real and imaginary parts in terms of *ber* and *bei* functions. He also replaces R^2/α^2 by its equivalent ν/n. The multiplying constant in (3) then has the simpler form $A/in\rho$, which is convenient for calculation, but does not show the relationship of the result to that for steady flow. It is more instructive to express the results in terms of modulus and phase, rather than in real and imaginary parts. Following the notation of McLachlan (1934) we write

$$J_0(xi^{3/2})=M_0(x)\exp i\theta_0(x)$$

and therefore (3) above may be written in the following form :

If the real part of Ae^{int} is $M\cos(nt-\phi)$ then

$$w=-\frac{AR^2}{\mu}\cdot\frac{M_0'}{\alpha^2}\sin(nt-\phi+\epsilon_0) \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where M_0' and ϵ_0 are defined as follows :

We have
$$J_0(\alpha i^{3/2}y)/J_0(\alpha i^{3/2})=\frac{M_0(\alpha y)}{M_0(\alpha)}\exp i\{\theta_0(\alpha y)-\theta_0(\alpha)\}.$$

Let $M_0(\alpha y)/M_0(\alpha)$ be written h_0, and write $\delta_0=\theta_0(\alpha)-\theta_0(\alpha y)$. Then

$$M_0'=\sqrt{(1+h_0^2-2h_0\cos\delta_0)}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

$$\tan\epsilon_0=h_0\sin\delta_0/(1-h_0\cos\delta_0). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

The rate of flow, Q, is given by

$$Q=2\pi R^2\int_0^1 wy\,dy \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

and integration of (3) gives at once

$$Q=\frac{\pi R^4}{\mu}\cdot\frac{A}{i^3\alpha^2}\left\{1-\frac{2J_1(\alpha i^{3/2})}{\alpha i^{3/2}J_0(\alpha i^{3/2})}\right\}e^{int} \qquad . \quad . \quad . \quad . \quad (8)$$

so that for a pressure-gradient $M\cos(nt-\phi)$ the rate of flow is

$$Q=\frac{\pi MR^4}{\mu}\cdot\frac{M_{10}'}{\alpha^2}\sin(nt-\phi+\epsilon_{10}) \qquad . \quad . \quad . \quad . \quad (9)$$

where, if $2M_1/\alpha M_0=h_{10}$, $\delta_{10}=(3\pi/4)-\theta_1(\alpha)+\theta_0(\alpha)$,

$$\left.\begin{aligned} M_{10}'&=\sqrt{(1+h_{10}^2-2h_{10}\cos\delta_{10})},\\ \tan\epsilon_{10}'&=h_{10}\sin\delta_{10}/(1-h_{10}\cos\delta_{10}).\end{aligned}\right\} \qquad . \quad . \quad . \quad (10)$$

The quantities M_{10}'/α^2 and ϵ_{10} have been tabulated by the Mathematics Division, National Physical Laboratory, for the range $0\leqslant\alpha\leqslant 10$, and are published (Womersley, in press).

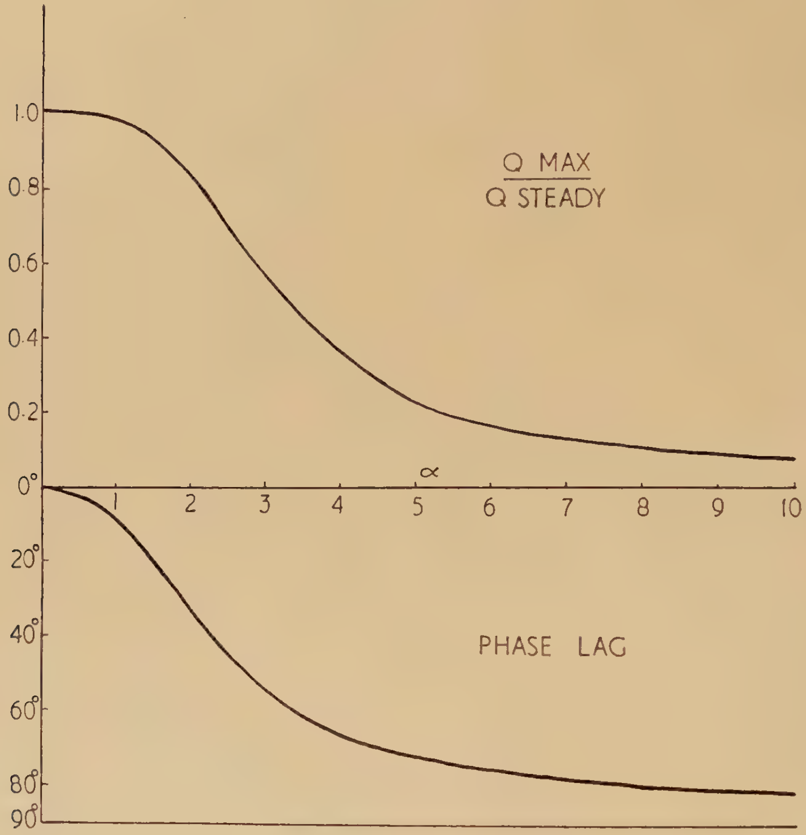
We make the following observations on this formula. First, the importance of the non-dimensional parameter α. This determines the character of the motion, and seems to vary very little (at corresponding points) in different animals. For example, in the femoral artery of the dog, which has a diameter of 3 mm, the pulse frequency being 3 per second, the value of α is 3·34. In the corresponding artery in the human thigh, the larger diameter is offset by a lower pulse-rate, and the value of α is about the same. The variation in phase-lag across the tube (somewhat analogous with the 'skin-effect' in electrical conductors) is considerable even at normal pulse-frequencies. Velocity-profiles for a range of values of α have been given by Lambossy (1952). The effect of α on the rate of flow is shown in fig. 1. The graph of phase-lag against α is self-explanatory. The graph of $Q_{\text{max}}/Q_{\text{steady}}$ gives the ratio of the modulus of the flow (i.e. its maximum value in either direction) to that of the Poiseuille flow for the same pressure-gradient. The values of α which are of interest in the study of the motion of the blood in the large arteries come on the steeply-falling parts of both curves. It is of interest to note also that both flow and phase-lag approach their asymptotic values quite slowly. For values of $\alpha>10$ there are simple asymptotic expressions for both M_{10}'/α^2 and ϵ_{10} which can be derived from the asymptotic expansions for the Bessel Functions. They are

$$M_{10}'/\alpha^2=1/\alpha^2-2/\alpha^3\sqrt{2}+1/\alpha^4,$$

$$\epsilon_{10}=2/\alpha\sqrt{2}+1/\alpha^2+19/24\sqrt{2}\alpha^2.$$

If the observed pressure-gradient is expressed as a Fourier Series, then the rate of flow can be calculated term by term, using the appropriate value of α for each harmonic, and the result compared with the observed rate of flow. This has been done, with fair agreement with observation (Womersley, in press) when one recalls the drastic nature of the assumption that the artery is a rigid tube, and that the formula contains no disposable constants.

Fig. 1



§ 3. The Equations of Motion of the Elastic Tube and the Derivation of the Pulse Velocity

In Lamb's paper (1898) " On the Velocity of Sound in a Tube as Influenced by the Elasticity of the Walls ", he considered a compressible inviscid fluid in a thin-walled elastic tube, took a wave-equation for the pressure inside and obtained a ' frequency equation ' for the velocity in terms of the dimensions of the tube and the elastic constants of the material. For a gas in a metal tube the correction is negligible, but he remarks " for the other extreme, e.g. water in an india-rubber tube, we have two values of c, $(hE/2a\rho_0)^{1/2}$ and $(B/\rho)^{1/2}$ ". In these expressions h is the thickness of the tube, a its radius, and ρ its density. E is Young's Modulus, and B is $E/(1-\sigma^2)$ where σ is Poisson's Ratio. ρ_0 is the density of the liquid.

Lamb adds that the former value of c is the same as the velocity of the pulse-wave as deduced theoretically by Résal (quoted by Korteweg in *Liouville's Journal* in 1876). It is apparent from Lamb's results that if the wavelength is long compared with the circumference of the tube, the velocity is independent of frequency. We show below that this does not hold if the velocity of the liquid is taken into account. Let R be the radius of the tube, h its thickness, ρ the density of the material, ρ_0 the density of the liquid. Take the axis of z along the tube, and let ζ be the displacement along the axis, ξ be the radial displacement. For the elastic constants take B to be $E/(1-\sigma^2)$ as above, where E is Young's modulus and σ Poisson's Ratio. Then the extensions are

$$\frac{\partial \zeta}{\partial z} \text{ along } z, \text{ and } \frac{\xi}{R} \text{ radial.}$$

The tensions are P along z, and Q radial, so that

$$P=Bh\left\{\frac{\partial \zeta}{\partial z}+\sigma\frac{\xi}{R}\right\}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

$$Q=Bh\left\{\frac{\xi}{R}+\sigma\frac{\partial \zeta}{\partial z}\right\}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

Lamb gives the equations of motion

$$\rho h\frac{\partial^2 \zeta}{\partial t^2}=\frac{\partial p}{\partial z}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

$$\rho h\frac{\partial^2 \xi}{\partial t^2}=p-\frac{Q}{R} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

where p is the excess pressure on the inner surface, but if the liquid is viscous there will be a surface traction on the inner surface of the tube due to viscous drag, and eqn. (13) must be modified. If w is the longitudinal component of the velocity of the liquid and u its radial component, the surface traction per unit area will be

$$\mu\left(\frac{\partial w}{\partial r}+\frac{\partial u}{\partial z}\right)_{r=R} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

where μ is the viscosity of the liquid. Making this modification to eqn. (13) and inserting the values of P and Q from (11) and (12) the equations of motion of the tube become

$$\frac{\partial^2 \zeta}{\partial t^2}=-\frac{\rho_0}{\rho}\frac{\nu}{hR}\left(\frac{\partial w}{\partial y}+R\frac{\partial u}{\partial z}\right)_{y=1}+\frac{B}{\rho}\left(\frac{\partial^2 \zeta}{\partial z^2}+\frac{\sigma}{R}\frac{\partial \xi}{\partial z}\right), \quad . \quad (16)$$

$$\frac{\partial^2 \xi}{\partial t^2}=\frac{p}{h\rho}-\frac{B}{\rho}\left(\frac{\sigma}{R}\frac{\partial \zeta}{\partial z}+\frac{\xi}{R^2}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

together with the boundary conditions for the motion of the liquid

$$u=\frac{\partial \xi}{\partial t} \text{ at } y=1, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

$$w=\frac{\partial \zeta}{\partial t} \text{ at } y=1. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

In the usual notation the equations of motion of the liquid are

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial r}+w\frac{\partial u}{\partial z}=-\frac{1}{\rho_0}\frac{\partial p}{\partial r}+\nu\left[\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}+\frac{\partial^2 u}{\partial z^2}-\frac{u}{r^2}\right], \quad . \quad (20)$$

$$\frac{\partial w}{\partial t}+u\frac{\partial w}{\partial r}+w\frac{\partial w}{\partial z}=-\frac{1}{\rho_0}\frac{\partial p}{\partial z}+\nu\left[\frac{\partial^2 w}{\partial r^2}+\frac{1}{r}\frac{\partial w}{\partial r}+\frac{\partial^2 w}{\partial z^2}\right] \quad . \quad . \quad . \quad (21)$$

together with the equation of continuity

$$\frac{\partial u}{\partial r}+\frac{u}{r}+\frac{\partial w}{\partial z}=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

If a trial solution is taken in which all the variables vary as $\exp\,[in\,(t-z/c)]$, we shall be concerned with the type of motion in which u/c, w/c and nR/c are all small. It is clear from the equation of continuity that u, the radial component of velocity, will be of order nR/c compared with w, the velocity along the tube. It follows that the non-linear terms are in general of order $1/c$ compared with the main linear terms, and we therefore neglect the inertia terms in the first approximation. Let

$$p=p_1\exp\,[in\,(t-z/c)],$$
$$u=u_1\exp\,[in\,(t-z/c)],$$
$$w=w_1\exp\,[in\,(t-z/c)].$$

Equations (20) and (21) reduce to

$$\frac{d^2u_1}{dy^2}+\frac{1}{y}\frac{du_1}{dy}+i^3\alpha^2u_1-\frac{u_1}{y^2}=\frac{1}{R}\cdot\frac{R^2}{\mu}\cdot\frac{\partial p_1}{\partial y}, \quad . \quad . \quad . \quad (23)$$

$$\frac{d^2w_1}{dy^2}+\frac{1}{y}\frac{dw_1}{dy}+i^3\alpha^2w_1 \qquad =-\frac{in}{c}\cdot\frac{R^2}{\mu}\cdot p_1 \quad . \quad . \quad . \quad (24)$$

where the terms in $\partial^2w/\partial z^2$, $\partial^2u/\partial z^2$ have been omitted, since n^2R^2/c^2 is small (see below).

The equation of continuity is

$$\frac{1}{y}\frac{d}{dy}(u\,.\,y)=\frac{inR}{c}w_1. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (25)$$

If now it is assumed that $p_1=AJ_0(ky)$, where k is to be determined, the equations of motion can be integrated to give

$$w_1=C_1\frac{J_0(\alpha i^{3/2}y)}{J_0(\alpha i^{3/2})}-\frac{inR^2}{c\mu}\frac{A_1}{i^3\alpha^2-k^2}J_0(ky), \quad . \quad . \quad . \quad (26)$$

$$u_1=C_2\frac{J_1(\alpha i^{3/2}y)}{J_0(\alpha i^{3/2})}-\frac{R}{\mu}\cdot k\cdot\frac{A_1}{i^3\alpha^2-k^2}J_1(ky), \quad . \quad . \quad . \quad (27)$$

where C_1 and C_2 are arbitrary constants, and the constant $J_0(\alpha i^{3/2})$ has been introduced (by analogy with the simple theory for the rigid tube) for convenience. If these values of u_1 and w_1 are inserted in the equation of continuity it should reduce to an identity. From (26) and (27)

$$\frac{inR}{c}w_1=\frac{inR}{c}C_1\frac{J_0(\alpha i^{3/2}y)}{J_0(\alpha i^{3/2})}-\frac{i^2n^2R^3}{c^2\mu}\cdot\frac{A}{i^3\alpha^2-k^2}J_0(ky), \quad . \quad (28)$$

$$\frac{1}{y}\frac{d}{dy}(u,\,y)=\alpha i^{3/2}\,C_2\frac{J_0(\alpha i^{3/2}y)}{J_0(\alpha i^{3/2})}-\frac{Rk^2}{\mu}\cdot\frac{A}{i^3\alpha^2-k^2}J_0(ky). \quad . \quad . \quad (29)$$

For the equation of continuity to be satisfied these must be identical, and therefore $k=inR/c$ and $C_2/C_1=inR/c\alpha i^{3/2}$, so that $J_0(ky)$ becomes $I_0(nR/c)$ and $J_1(ky)$ becomes $iI_1(nR/c)$.

At the highest frequencies likely to be of interest in the pulse-wave, n^2R^2/c^2 is small compared with α^2. As an example, for the sixth harmonic of the pulse-frequency of the dog, $n^2R^2/c^2 \sim 6\times 10^{-4}$, and the ratio of this to the corresponding value of α^2 is about 9×10^{-6}. It is therefore justifiable to replace $i^3\alpha^2-k^2$ in (26) and (27) by $i^3\alpha^2$, and in view of the smallness of nR/c to use the approximations $I_0(nR/c)=1$ and $I_1(nRy/c)=nRy/2c$. This is the same degree of approximation that is implicit in omitting the terms $\partial^2 w/\partial z^2$ and $\partial^2 u/\partial x^2$ from the equations of motion. Inserting these approximations in the expressions for the velocity-components,

$$w_1=C_1\frac{J_0(\alpha i^{3/2}y)}{J_0(\alpha i^{3/2})}+\frac{A_1}{\rho_0 c}, \qquad (30)$$

$$u_1=\frac{inR}{2c}\left[C_1\frac{2J_1(\alpha i^{3/2}y)}{\alpha i^{3/2}J_0(\alpha i^{3/2})}+y\,.\,\frac{A_1}{\rho_0 c}\right]. \qquad (31)$$

At the inner surface of the tube, i.e. when $y=1$,

$$w_1=C_1+\frac{A_1}{\rho_0 c}, \qquad (32)$$

$$u_1=\frac{1}{2}\,.\,\frac{inR}{c}\,.\,C_1\,.\,F_{10}(\alpha)+\frac{1}{2}\frac{inR}{c}\,.\,\frac{A_1}{\rho_0 c} \qquad (33)$$

where $F_{10}(\alpha)$ has been written for $2J_1(\alpha i^{3/2})/\alpha i^{3/2}J_0(\alpha i^{3/2})$, the function of α which appears in the expression for the rate of flow in the simple theory. The value of $\partial w_1/\partial y$ at $y=1$ is also needed, to substitute in the equations of motion of the tube. Making the same approximations, this is

$$-\frac{C_1}{2}\,.\,i^3\alpha^2\,F_{10}(\alpha)+\frac{1}{2}\frac{n^2R^2}{c^2}\,.\,\frac{A_1}{\rho_0 c}. \qquad (34)$$

If it is now assumed that

$$\xi=D_1\exp\,[in\,(t-2/c)],$$
$$\zeta=E_1\exp\,[in\,(t-2/c)],$$

where D_1 and E_1 are arbitrary constants, the boundary conditions for u_1 and w_1 become

$$in\,E_1=C_1+\frac{A_1}{\rho_0 c}, \qquad (35)$$

$$in\,D_1=\frac{1}{2}\frac{inR}{c}\left\{F_{10}(\alpha)\,.\,C_1+\frac{A_1}{\rho_0 c}\right\} \qquad (36)$$

and the equations of motion of the tube become

$$-n^2D_1=\frac{A_1}{h\rho}-\frac{B}{\rho}\left\{\frac{\sigma}{R}\left(-\frac{inE_1}{c}\right)+\frac{D_1}{R^2}\right\}, \qquad (37)$$

$$\begin{aligned}-n^2E_1=&-\frac{\rho_0}{\rho}\,.\,\frac{\nu}{hR}\left\{-\frac{C_1}{2}\,.\,i^3\alpha^2\,F_{10}(\alpha)+\frac{1}{2}\frac{A_1}{\rho_0 c}\,.\,\frac{n^2R^2}{c^2}\right\}\\&+\frac{B}{\rho}\left\{-\frac{n^2}{c^2}E_1+\frac{\sigma}{R}\left(-\frac{in}{c}\right)D_1\right\}.\end{aligned} \qquad (38)$$

In (38) the term in $\partial u/\partial z$ in (16) has been omitted, since it is of order n^2R^2/c^2.

Equations (35)–(38) are four homogeneous equations in the arbitrary constants A_1, C_1, D_1, E_1. Eliminating them will give a 'frequency equation' which will determine the wave-velocity c in terms of the elastic properties of the tube and the non-dimensional parameter α. The result of the elimination is

$$0=\begin{vmatrix} 1/\rho_0 c & 1 & 0 & -in \\ inR/2\rho_0 c^2 & inR/2c & -in & 0 \\ 1/h\rho & 0 & n^2-B/\rho R^2 & in\sigma B/cR\rho \\ \frac{1}{2}\cdot\frac{\rho_0}{\rho}\cdot\frac{\nu}{hR}\cdot\frac{1}{\rho_0 c}\cdot\frac{n^2R^2}{c^2} & -\frac{i}{2}\cdot\frac{\rho_0}{\rho}\frac{\nu\alpha^2}{hR}F_{10} & -\frac{in}{c}\cdot\frac{\sigma}{R}\cdot\frac{B}{\rho} & n^2\left(1-\frac{B}{\rho c^2}\right) \end{vmatrix}$$

The first step in the evaluation of this determinant is to make all the elements non-dimensional. The term $-i\rho_0\nu\alpha^2 F_{10}/2\rho hR$ is replaced by its equivalent $-in\rho_0 RF_{10}/2\rho h$, since $\nu\alpha^2=nR^2$. If now the first column is divided by $\rho_0 c$, the second column by inR/c, fourth column by $-in$, and the fourth row by in, the determinant becomes

$$\begin{vmatrix} 1 & 1 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2}F_{10} & -1 & 0 \\ R\rho_0/h\rho & 0 & -B/\rho c^2 & -\sigma B/\rho c^2 \\ 0 & -F_{10}R\rho_0/2h\rho & -\sigma B/\rho c^2 & 1-B/\rho c^2 \end{vmatrix}$$

At this stage the following further approximations have been made. In the third row, third term, n^2R^2/c^2 has been neglected in comparison with $B/\rho c^2$. This is justified, since $(B/\rho)^{1/2}$ is the velocity of shear-waves in the material of the tube, and this will be greater than the pulse-velocity, so that $B/\rho c^2>1$, and n^2R^2/c^2 can be neglected in comparison with it. The first term of the fourth row is of order $n^2R^2/c^2\alpha^2$ and can therefore be neglected. If now we write $k=h\rho/R\rho_0$, $x=kB/\rho_0{}^2$, and reduce the order of the determinant by replacing the second row by (Row 1$-2\times$Row 2) and the third row by (Row 1$-$Row 3) the equation becomes

$$0=\begin{vmatrix} 1-F_{10} & 2 & 1 \\ 1 & x & 1+\sigma x \\ -\frac{1}{2}F_{10} & -\sigma x & (k-x) \end{vmatrix}$$

which reduces to

$$0=(1-\sigma^2)(1-F_{10})x^2-x\{2+k(1-F_{10})+F_{10}(\tfrac{1}{2}-2\sigma)\}+F_{10}+2k. \quad . \quad (39)$$

In this equation $k\ll 1$ since the tube is assumed to be thin-walled, and in the application to the arteries $\rho=\rho_0$.

When $\alpha\rightarrow\infty$, i.e. as the viscosity of the liquid $\rightarrow 0$, $F_{10}\rightarrow 0$ and the roots of this equation tend to equality with those of the corresponding equation which can be deduced from Lamb's equation for the compressible fluid.

An equation which is essentially the same as (39) has been given by Morgan and Kiely (1954) in a somewhat different notation. They give two approximate solutions, one for ‘ small viscosity ’, by using the asymptotic expansions of the Bessel Functions, and the other for ‘ large viscosity ’, by using the power series, but do not seem to have appreciated the dependence of the wave-velocity on the single non-dimensional parameter, α.

The roots of (39) are given by

$$(1-\sigma^2)x=G\pm\sqrt{\{G^2-(1-\sigma^2)H\}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (40)$$

where

$$G=\frac{1+\frac{1}{4}-\sigma}{1-F_{10}}+(k/2+\sigma-\tfrac{1}{4}), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (41)$$

$$H=(1+2k)/(1-F_{10})-1. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (42)$$

In terms of the notation used in (9) above, $1/(1-F_{10})=\exp(-i\epsilon)/M_{10}'$ so that the quantities required to compute the roots are already available. Since $F_{10}(\alpha)$ is complex, x is always complex, and the motion is either damped or unstable. The sign of arg x is determined by that of arg G, and since arg G is always negative, the motion is damped. If we write

$$\{(1-\sigma^2)x/2\}^{1/2}=X-iY \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (43)$$

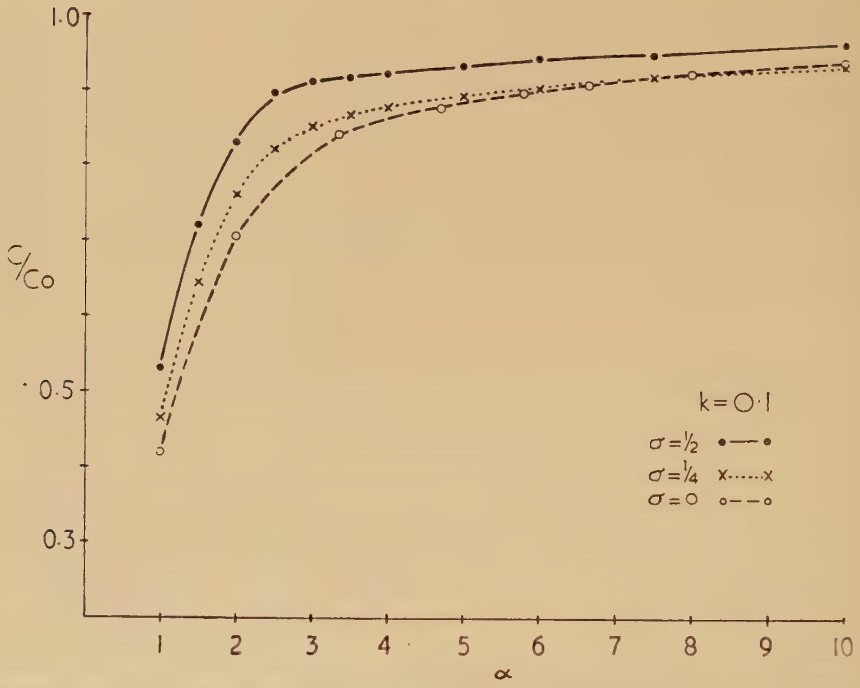
and denote by c_0 the velocity for the perfect fluid, i.e., write $c_0=(hE/2R\rho_0)^{1/2}$, then if c_1 is the wave-velocity, $c_0/c_1=X$ and over a distance of one wavelength the amplitude will be reduced in the ratio $\exp(-2\pi Y/X)$. The quantities of practical interest are, therefore, $1/X$, which gives the ratio of the velocity to that for the perfect fluid, and $\exp(-2\pi Y/X)$. It is hoped to publish full tables of these, together with other quantities required for calculating the longitudinal velocity and rate of flow of the liquid, in a later communication. Table 1 gives their values, together with the reduction in amplitude over a distance of 10 cm, for the first four harmonics of the pulse of the dog, in the femoral artery. They have been calculated for $k=0{\cdot}1$, i.e., for a wall-thickness of one-tenth of the radius, for $\sigma=\frac{1}{2}$ and $\sigma=0$. In this table, $f_1=\exp(-2\pi Y/X)$ and $f_2=\exp(-2\pi Yz/X\lambda)$, with $z=10$ cm.

Table 1

	$\sigma=\frac{1}{2}$			$\sigma=0$		
α	c/c_0	f_1	f_2	c/c_0	f_1	f_2
3·34	0·914	0·274	0·946	0·842	0·132	0·917
4·72	0·924	0·472	0·938	0·876	0·294	0·900
5·78	0·936	0·565	0·929	0·894	0·381	0·883
6·67	0·942	0·636	0·925	0·906	0·442	0·870

The amount of damping seems rather large. A diminution in amplitude of 5–6% over so short a length of 10 cm one would expect to have been observed and remarked upon, but until more accurate observations are available it is not possible to say with certainty that this degree of damping is greater than that which exists in the arterial system. In practical observation it might well be masked by the change in shape of the pulse as it travels. Figure 3, which shows the relationship between α and $\exp(-2\pi Y/X)$, illustrates how much greater the damping is for the smaller values of σ. Figure 2, which shows $1/X = c/c_0$ against α, demonstrates that

Fig. 2



for values of α greater than 3, which are those of greatest practical interest, the change in velocity with α is quite small. It is somewhat larger the smaller is the value of σ. Since the wave-velocity varies with frequency, energy will be transmitted with the group-velocity, and it is the group-velocity that will be observed when the pulse-velocity is measured. If a wave-motion is described by the form $\exp[i(nt-mz)]$, the group velocity c_g is given by $c_g = dn/dm$. In terms of α and X this may be written

$$\frac{c_0}{c_g} = \frac{c_0}{c_1}\left(1 - \frac{\alpha}{2X}\frac{dX}{d\alpha}\right). \qquad (44)$$

The analytical form for $dX/d\alpha$ is unsuitable for computation. Figure 4 shows a graph of $\log X$ against $\log \alpha$, and since

$$\frac{\alpha}{X}\frac{dX}{d\alpha} = \frac{d(\log X)}{d(\log \alpha)} \qquad (45)$$

Fig. 3

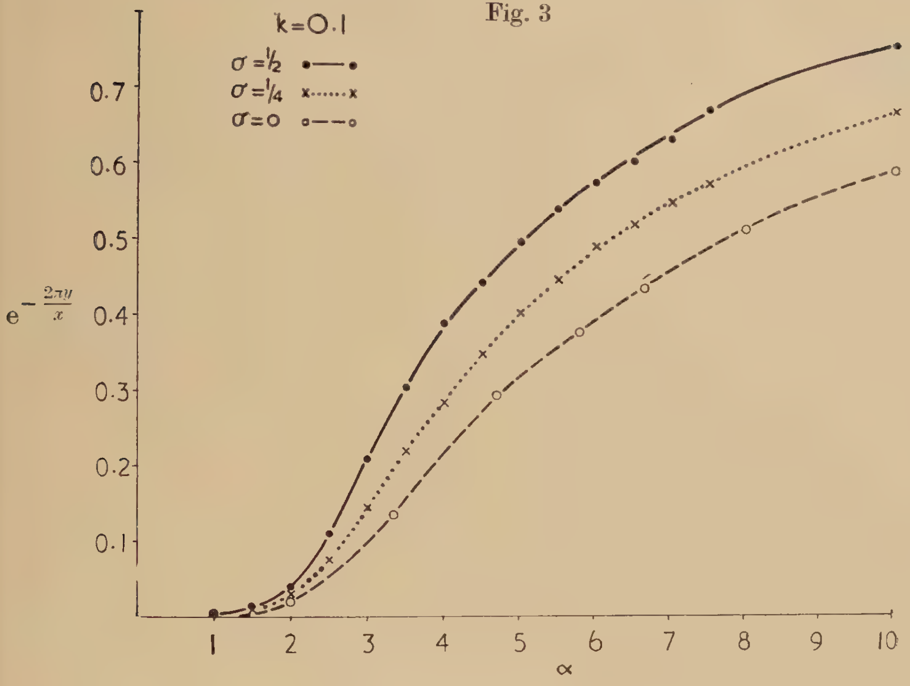
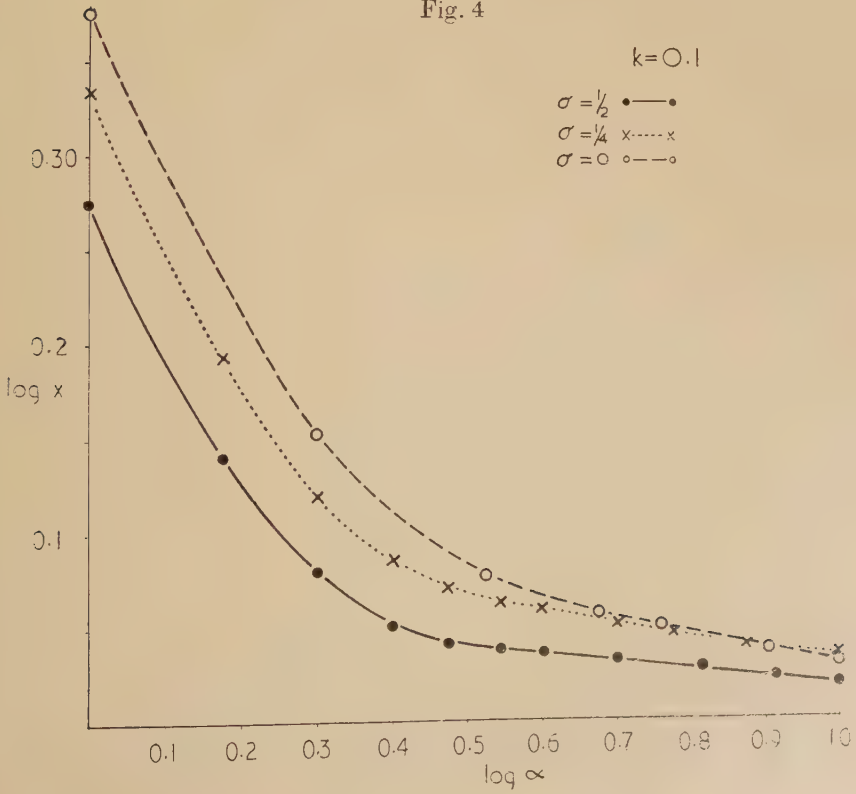

k=0.1
σ=½ •—•
σ=¼ x·····x
σ=0 o—–o
$e^{-\frac{2\pi y}{x}}$
0.7
0.6
0.5
0.4
0.3
0.2
0.1
1
2
3
4
5
6
7
8
9
10
∝

Fig. 4


k=0.1
σ=½ •—•
σ=¼ x·····x
σ=0 o—–o
0.30
0.2
log x
0.1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
log ∝

it is possible from this to estimate its magnitude. It will be seen from the graph that for those values of α which apply to the femoral artery $(3<\alpha<7)$, $d(\log X)/d(\log \alpha)$ is about $-0{\cdot}045$, so that for these conditions c_g/c_1 is about 0·98, so that the difference is only 2%, and over the range $\alpha=3$ to $\alpha=4$ is certainly never more than $2\frac{1}{2}$%. Until accurate measurements of pulse-velocity are made over short lengths of artery, this effect is not likely to be worth taking into account.

§ 4. THE MOTION OF THE LIQUID

In this section the details of the motion of the liquid at a particular value of z, i.e., over a short length of artery, will be studied. If the origin of z is taken at the mid-point of this short length, the longitudinal velocity will be

$$w=\left\{\frac{A_1}{\rho_0 c}+C_1\frac{J_0(\alpha i^{3/2}y)}{J_0(\alpha i^{3/2})}\right\}e^{int} \quad . \quad . \quad . \quad . \quad . \quad . \quad (46)$$

where the value of C_1 is to be determined from the boundary-conditions. From (35), (36) and (37), D_1 and E_1 can be eliminated. If, in the resulting equation, η is written for the ratio of C_1 to $A_1/\rho_0 c$, the value of η is given by

$$\eta=\frac{2}{x(F_{10}-2\sigma)}-\frac{(1-2\sigma)}{F_{10}-2\sigma} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (47)$$

where x is the root of eqn. (39) and then

$$w=\frac{A_1}{\rho_0 c}\left\{1+\eta\frac{J_0(\alpha i^{3/2}y)}{J_0(\alpha i^{3/2})}\right\}e^{int}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (48)$$

The first point to be noted is that for a given pressure-function, the pressure-gradient is inversely proportional to the wave-velocity, so that the velocity of the liquid, and therefore the amount of flow, will be proportionately greater in that which is more elastic. The constant, $A_1/\rho_0 c$ may be written

$$\left(-\frac{in}{c}\right).\frac{1}{\rho_0}.\left(-\frac{1}{in}\right)A_1$$

and since $\alpha^2=nR^2/\nu$ this may be written

$$A_1/\rho_0 c=A_1{}'R^2/i^3\alpha^2\mu \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (49)$$

where $A_1{}'$ is the quantity that will be given by Fourier Analysis of an observed pressure-gradient. This establishes the identity of the constant in (49) with that in the simple theory for the rigid tube. The average velocity across the tube is

$$\bar{w}=\frac{A_1}{\rho_0 c}\{1+\eta F_{10}\}e^{int}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (50)$$

This can be put in the form

$$\bar{w}=\frac{A_1{}'R^2}{\mu}\frac{M_{10}{}''}{\alpha^2}\sin(nt-\phi+\epsilon_{10}{}'') \quad . \quad . \quad . \quad . \quad (51)$$

for a pressure-gradient $M\cos(nt-\phi)$, and the values of $M_{10}{}''$ and $\epsilon_{10}{}''$ compared with those of $M_{10}{}'$ and $\epsilon_{10}{}'$ for the rigid tube. For $\sigma=\frac{1}{2}$ and $\sigma=0$ the expressions for η and ηF_{10} are particularly simple.

When $\sigma=\frac{1}{2}$, $\eta=-2/x(1-F_{10})$,

$$1+\eta F_{10}=1-2F_{10}/x(1-F_{10}). \qquad (52)$$

When $\sigma=0$, $\eta=(1-2/x)/F_{10}$,

$$1+\eta F_{10}=2/x. \qquad (53)$$

Using these formulae, and the corresponding one for $\sigma=\frac{1}{4}$, the values of M_{10}'' and ϵ_{10}'' have been calculated for $\sigma=\frac{1}{2}, \frac{1}{4}, 0$, for the first four harmonics in the femoral artery. The results are given in table 2. It will be seen that if the wall is 'perfectly elastic', i.e., suffers no volume change in stretching, the average velocity is some 10% more than that for a rigid tube under the same pressure-gradient. At the other extreme, if the wall is composed of a material which does not form a 'waist' on stretching, the velocity is slightly less than that for a rigid tube. The phase-lag for $\sigma=\frac{1}{2}$ is greater than that for the rigid tube, but for the smaller values of σ is less, and this difference in phase-lag becomes smaller as α increases.

Table 2

	$\sigma=\frac{1}{2}$		$\sigma=\frac{1}{4}$		$\sigma=0$		Rigid	
α	M_{10}''	ϵ_{10}''	M_{10}''	ϵ_{10}''	M_{10}''	ϵ_{10}''	M_{10}'	ϵ_{10}'
3·34	0·7373	29·26°	0·7071	32·11°	0·6379	35·62°	0·6551	30·98°
4·72	0·7945	17·86°	0·7767	20·29°	0·7378	22·38°	0·7436	19·95°
5·78	0·8275	13·71°	0·8117	15·77°	0·7807	17·48°	0·7839	15·82°
6·67	0·8471	11·59°	0·8279	13·58°	0·8076	14·80°	0·8096	13·49°

Table 3

	$\sigma=\frac{1}{2}$		$\sigma=0$	
α	Amplitude ratio	Phase-difference	Amplitude ratio	Phase-difference
3·34	0·122	74·0°	0·527	26·2°
4·72	0·166	52·9°	0·226	16·2°
5·78	0·190	38·0°	0·145	12·6°
6·67	0·257	28·6°	0·104	10·5°

It is also of interest to calculate the longitudinal velocity at the wall. This is

$$\frac{A_1}{\rho_0 c}(1+\eta)\,e^{int}$$

so that its ratio to the average velocity is $(1+\eta)/(1+\eta F_{10})$. If the modulus and phase of this are calculated, they will demonstrate the ratio of the magnitude of the velocity at the wall to that of the average velocity, and the phase-difference between the two. Table 3 shows these quantities, for $\sigma=\frac{1}{2}$, and $\sigma=0$, for the same values of α as in tables 1 and 2.

These figures emphasize the critical role of Poisson's Ratio in determining the details of the motion. The phase-differences in the above table are shown with their correct sign in relation to the amplitude ratio, and therefore show that the velocity at the wall 'leads' the average velocity. The most striking point about this effect is its magnitude, which is greater than might be expected. It is not possible to estimate it from the experimental results available, but it would seem that the attempt to find experimental means of measuring the effect would be worth while, since it would form a critical test of the theory, and would throw light on the elasic properties of the arterial wall under dynamic conditions.

The formula for the rate of flow might at first sight be written

$$Q=\frac{A_1}{\rho_0 c}\pi R^2(1+\eta F_{10})e^{int} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (54)$$

but this is approximate only, since at any time the value of the radius is not R, but $R+\xi$, and ξ varies with the time. A better approximation should be given, therefore, by

$$Q=\frac{A_1}{\rho_0 c}\cdot\pi R^2\left(1+\frac{2\xi}{R}\right)(1+\eta F_{10})e^{int}. \quad . \quad . \quad . \quad . \quad (55)$$

Now $2\xi/R=\bar{w}/c$, a useful formula for rough use in estimating whether the arterial expansion observed in an experiment is consistent with the observations of pressure-gradient and average velocity. Even (55) is not fully corrected for variations in radius. In the original equations, R has been regarded as constant from the point of view of motion of the liquid. In fact the boundary of the liquid is at $R+\xi$. If we replace R by $R+\xi$ in (20) and assume that w is still essentially a function of y (i.e. that the lines of laminar flow expand and contract with the artery) (20) may be written

$$\frac{\partial^2 w}{\partial y^2}+\frac{1}{y}\frac{\partial w}{\partial y}-\frac{R^2}{\nu}\left(1+\frac{2\xi}{R}\right)\frac{\partial w}{\partial t}=\left(1-\frac{2\xi}{R}\right)\frac{R^2}{\mu}\frac{\partial p}{\partial z} \quad . \quad . \quad (56)$$

the inertia terms and the term in $\partial^2 w/\partial z^2$ being omitted.

If we seek a solution which is of the same form as that for a constant radius, we can imagine $\partial p/\partial z$ represented by a Fourier Series in $n(t-z/c)$, and corresponding Fourier Series for w and $2\xi/R$ (which can be represented in terms of $\bar{w}$) substituted in the equation. The products of the Fourier Series can be multiplied out, and a set of equations for the velocity components obtained by collecting up corresponding terms. The wave-velocity is not the same at all frequencies, however, and on multiplying two periodic terms together there will be exponential factors in z 'left over' as it were, which would disappear (being equal to unity) in a system with constant wave-velocity. A typical term is

$$\exp\{in(t-z/c_1)\}\times\exp\{in(t-z/c_1)\}$$

which has to be compared with $\exp\{2in(t-z/c_2)\}$ where c_1 is the wave-velocity for a frequency $n/2\pi$, and c_2 the wave-velocity for double this frequency. The question is, therefore, how far does $\exp\{2in(1/c_1-1/c_2)z\}$ differ from unity ? A glance at the graph of c/c_0 against α shows that for $\alpha>3$ the variation in c is small, and in fact for the first four harmonics of the pulse of the dog, all these ‘ stray ’ exponential terms have wavelengths of ten metres or more.

In the next section, therefore, a simple perturbation method, along the lines described, is used to solve (56). It is admitted that this is not necessarily the most important correction to be made to the linear solution given above. But it is the simplest to do, and provides a ‘ lead-in ’ to the more important correction for the inertia terms. And though at first sight less important in principle than the inertia-term correction, it may well be equally important in magnitude, for it may cope with fairly large arterial distentions, such as occur near the heart, without the mathematical difficulties that arise in the analysis of finite strain.

§ 5. FIRST-ORDER CORRECTION FOR VARIATION IN DIAMETER

A solution is sought of eqn. (56) when the pressure-gradient is a Fourier Series of four harmonics, together with a constant term which will be assumed to give a Poiseuille flow, the static expansion of the tube (which would give a tapering effect) being neglected. This is justified, since this constant term is small, being less than one-eighth of the largest oscillatory terms. The detailed solution will be developed for two harmonics only, since this illustrates the method adequately without wearisome repetition, but the numerical example quoted has been worked out for a series with four harmonics. Since (56) is non-linear, we may no longer write the pressure-gradient as

$$A_0+A_1\,e^{int}+A_2\,e^{2int}+\dots$$

and take the real part, or half the interaction terms will be lost. It is necessary to start from the pressure-gradient in real form, and write down its exact complex equivalent. Assume that

$$\partial p/\partial z=M_0+M_1\cos(nt+\phi_1)+M_2\cos(2nt+\phi_2). \quad . \quad . \quad (57)$$

If we define A_0, A_1, A_2 by

$$A_0=M_0,\quad A_1=\tfrac{1}{2}M_1\exp(i\phi_1),\quad A_2=\tfrac{1}{2}M_2\exp(i\phi_2)$$

then

$$\partial p/\partial z=A_0+A_1\,e^{int}+A_1{}^*\,e^{-int}+A_2\,e^{2int}+A_2{}^*\,e^{-2int} \quad . \quad . \quad (58)$$

using a standard notation for conjugate complex quantities. Taking the value of $\overline{w}$ already obtained as an adequate approximation to use in substituting for $2\xi/R$ in eqn. (56) we have

$$1+2\xi/R=1+\frac{1}{c_0}(C_1\,e^{int}+C_1{}^*\,e^{-int}+C_2\,e^{2int}+C_2{}^*\,e^{-2int}) \qquad (59)$$

in which

$$C_1=(X_1-iY_1)\frac{A_1R^2}{\mu i^3\alpha_1{}^2}(1+\eta_1F_{10}), \quad . \quad . \quad . \quad . \quad . \quad (60)$$

$$C_2=(X_2-iY_2)\frac{A_2R^2}{\mu i^3\alpha_2{}^2}(1+\eta_2F_{20}). \quad . \quad . \quad . \quad . \quad . \quad (61)$$

In (60) and (61) the suffixes indicate that the values of α and quantities depending upon it appropriate to the harmonic. If (58) and (59) are now inserted in the right-hand side of (56) we have, omitting the constant factor R^2/μ,

$$\begin{aligned}(1+2\xi/R)\frac{\partial p}{\partial z}=&A_0+A_1\,\mathrm{e}^{int}+A_1{}^*\,\mathrm{e}^{-int}+A_2\,\mathrm{e}^{2int}+A_2{}^*\,\mathrm{e}^{-2int}\\ &+\frac{A_0}{c_0}(C_1\,\mathrm{e}^{int}+C_1{}^*\,\mathrm{e}^{-int}+C_2\,\mathrm{e}^{2int}+C_2{}^*\,\mathrm{e}^{-int})\\ &+\frac{1}{c_0}(A_1\,\mathrm{e}^{int}+A_1{}^*\,\mathrm{e}^{-int})(C_1\,\mathrm{e}^{int}+C_1{}^*\,\mathrm{e}^{-int}+C_2\,\mathrm{e}^{2int}\\ &+C_2{}^*\,\mathrm{e}^{-int})\\ &+\frac{1}{c_0}(A_2\,\mathrm{e}^{2int}+A_2{}^*\,\mathrm{e}^{-2int})(C_1\,\mathrm{e}^{int}+C_1{}^*\,\mathrm{e}^{-int}+C_2\,\mathrm{e}^{2int}\\ &+C_2{}^*\,\mathrm{e}^{-2int}). \quad \ldots \ldots \quad (62)\end{aligned}$$

Now consider the term in $1+\xi/2R$ on the left-hand side of (56). Assuming

$$w=w_0+w_1\,\mathrm{e}^{int}+w_1{}^*\,\mathrm{e}^{-int}+w_2\,\mathrm{e}^{2int}+w_2{}^*\,\mathrm{e}^{-2int}$$

this becomes

$$\begin{aligned}\frac{R^2}{\nu}\left(1+\frac{2\xi}{R}\right)\frac{\partial w}{\partial t}=&i^3\alpha^2(w_1\,\mathrm{e}^{int}-w_1{}^*\,\mathrm{e}^{-int})+2i^3\alpha^2(w_2\,\mathrm{e}^{2int}+w_2{}^*\,\mathrm{e}^{-2int})\\ &+\frac{i^3\alpha^2}{c_0}(w_1\,\mathrm{e}^{int}-w_1{}^*\,\mathrm{e}^{-int})(C_1\,\mathrm{e}^{int}+C_1\,\mathrm{e}^{-int}+C_2\,\mathrm{e}^{2int}\\ &+C_2{}^*\,\mathrm{e}^{-2int})\\ &+\frac{2i^3\alpha^2}{c_0}(w_2\mathrm{e}^{2int}-w_2{}^*\,\mathrm{e}^{-2int})(C_1\,\mathrm{e}^{int}+C_1{}^*\,\mathrm{e}^{-int}+C_2\,\mathrm{e}^{2int}\\ &+C_2{}^*\,\mathrm{e}^{-2int}).\end{aligned}$$

If now (60) and (62) are inserted in (56), and corresponding powers of e^{int} are collected, a set of equations for w_0, w_1, $w_1{}^*$, w_2, $w_2{}^*$ can be written down. The terms independent of e^{int} give the equation

$$\begin{aligned}\frac{d^2w_0}{dy^2}+\frac{1}{y}\frac{dw_0}{dy}=&\frac{A_0R^2}{\mu}+\frac{1}{c_0}\frac{R^2}{\mu}(A_1C_1{}^*+A_1{}^*C_1+A_2C_2{}^*+A_2{}^*C_2)\\ &+\frac{i\alpha^2}{c_0}\{C_1{}^*w_1-C_1w_1{}^*+2(C_2{}^*w_2-C_2w_2{}^*)\} \quad \ldots \quad (63)\end{aligned}$$

and from the terms in e^{int}, e^{2int} give

$$\begin{aligned}\frac{d^2w_1}{dy^2}+\frac{1}{y}\frac{dw_1}{dy}+i^3\alpha^2w_1=&\frac{A_1R^2}{\mu}+\frac{R^2}{\mu c_0}(A_0C_1+A_2C_1{}^*+A_1{}^*C_2)\\ &+\frac{i\alpha^2}{c_0}(-C_2w_1{}^*+2C_1{}^*w_2), \quad \ldots \quad (64)\end{aligned}$$

$$\frac{d^2w_2}{dy^2}+\frac{1}{y}\frac{dw_2}{dy}+2i^3\alpha^2w_2=\frac{A_2R^2}{\mu}+\frac{R^2}{\mu c_0}(A_0C_2+A_1C_1)+\frac{i\alpha^2}{c_0}C_1w_1. \quad \ldots \quad (65)$$

An approximate solution, correct to order $1/c_0$ can be obtained by inserting the known forms for w_1 and $w_1{}^*$, w_2 and $w_2{}^*$ in the right-hand sides of these equations. If this is done the integration can be carried through and the result expressed in terms of functions already known. The equation for w_0 becomes

$$\frac{d^2w_0}{dy^2}+\frac{1}{y}\frac{dw_0}{dy}=\frac{A_0R^2}{\mu}+\frac{1}{c_0}(A_1C_1^*+A_1^*C_1+\ldots)$$
$$+\frac{i\alpha^2}{c_0}\left[\frac{C_1^*R^2}{\mu i^3\alpha^2}\left\{1+\eta_1\frac{J_0(\alpha i^{3/2}y)}{J_0(\alpha i^{3/2})}\right\}+\ldots\right].$$

It will be seen that, with the exception of the main term, the constant terms cancel, so that this becomes

$$\frac{d^2w_0}{dy^2}+\frac{1}{y}\frac{dw_0}{dy}=\frac{A_0R^2}{\mu}-\frac{R^2}{c_0\mu}\left\{C_1^*A_1\eta_1\frac{J_0(\alpha i^{3/2}y)}{J_0(\alpha i^{3/2})}+C_1A_1^*\eta_1^*\frac{J_0(\alpha i^{-3/2}y)}{J_0(\alpha i^{-3/2})}\right\}. \quad . \quad . \quad . \quad (66)$$

In this equation only the first pair of terms has been retained, the other pair being identical in form.

Since for any Bessel Function of zero order

$$\left(\frac{d^2}{dy^2}+\frac{1}{y}\frac{d}{dy}\right)J_0(ky)=-k^2J_0(ky)$$

the solution of (66) which is zero at $y=1$ is

$$w_0=-\frac{A_0R^2}{\mu}(1-y^2)-\frac{1}{c_0}\frac{R^2}{\mu}\eta_1\frac{C_1^*A_1}{i^3\alpha^2}\left\{1-\frac{J_0(\alpha i^{3/2}y)}{J_0(\alpha i^{3/2})}\right\}$$
$$-\frac{1}{c_0}\frac{R^2}{\mu}\eta_1^*\frac{C_1A_1^*}{i^{-3}\alpha^2}\left\{1-\frac{J_0(\alpha i^{-3/2}y)}{J_0(\alpha i^{-3/2})}\right\}. \quad . \quad . \quad (67)$$

If in this we replace

$$\left\{1-\frac{J_0(\alpha i^{3/2}y)}{J_0(\alpha i^{3/2})}\right\} \quad \text{by} \quad 1-F_{10}$$

and

$$\left\{1-\frac{J_0(\alpha i^{-3/2}y)}{J_0(\alpha i^{-3/2})}\right\} \quad \text{by} \quad 1-F_{10}^*$$

we shall have the expression for the average velocity.

Inserting the vaules of C_1 and C_1^*, and writing

$$X_1+iY_1=N_1\exp(i\delta_1), \qquad -\eta_1=k\exp(i\gamma_1)$$

the contribution to the average velocity from this pair of correction terms is

$$M_{10}'\,.\,M_{10}''\,.\,\frac{M_1^2R^4}{c_0\mu^2\alpha^4}\cos(\epsilon_{10}'-\epsilon_{10}''+\delta_1+\gamma_1) \quad . \quad . \quad . \quad (68)$$

and there will be another contribution from the other pair of terms, identical in form, but with values appropriate to the second harmonic. This correction is in opposition to the main term. When $\sigma=\frac{1}{2}$ this expression can be simplified, for then $k_1=\frac{3}{4}N_1^2M_{10}'$ and $\gamma_1=2\delta_1-\epsilon_{10}'$.

Substituting these values, and replacing $R^2/\mu\alpha^2$ by its equivalent $1/\rho n$. (68) becomes, when $\sigma=\frac{1}{2}$,

$$\frac{1}{c_0}\cdot\frac{3}{4}\cdot\frac{M_{10}''}{\rho^2 n^2}\frac{M_1^{\,2}}{N_1}\cos(3\delta_1-\epsilon_{10}''). \quad . \quad . \quad . \quad . \quad . \quad . \qquad (69)$$

In the equation for the first harmonic, the constant terms cancel each other in the same way, and the equation reduces to

$$\frac{d^2w_1}{dy^2}+\frac{1}{y}\frac{dw_1}{dy}+i^3\alpha^2w_1=\frac{R^2}{\mu}A_1+\frac{1}{c_0}\frac{R^2}{\mu}A_0C_1-\frac{1}{c_0}\frac{R^2}{\mu}C_2A_1{}^*\eta_1{}^*\frac{J_0(\alpha_1 i^{-3/2}y)}{J_0(\alpha_1 i^{-3/2})}$$
$$-\frac{1}{c_0}\frac{R^2}{\mu}C_1{}^*A_2\eta_2\frac{J_0(\alpha_2 i^{3/2}y)}{J_0(\alpha_2 i^{3/2})}. \quad . \quad . \quad . \quad . \quad . \quad . \qquad (70)$$

The solution of this equation can be written down at once by repeated application of the well-known result :

If
$$\frac{d^2w}{dy^2}+\frac{1}{y}\frac{dw}{dy}+k^2w=A_0\frac{J_0(ly)}{J_0(l)}$$

then the solution of this which is zero at $y=1$ is

$$w=\frac{A}{k^2-l^2}\left\{\frac{J_0(ly)}{J_0(l)}-\frac{J_0(ky)}{J_0(k)}\right\}. \quad k\neq l. \quad . \quad . \quad . \quad . \quad . \qquad (71)$$

In imposing the condition that all the correction terms vanish at $y=1$, a further approximation is being made. Physically, it enforces the condition that the motion of the wall is due to the main terms only and the correction terms have no effect. Since the correction terms are small, this approximation may be adequate. For complete consistency, the arbitrary constants in the expressions for average velocity which are substituted in the equation should be left ‘ floating ’ and the fully corrected solution substituted back in the equations of motion of the tube. The ‘ frequency-equation ’ would then be non-linear, and the pulse-velocity would depend on the particular form of the pressure function. The same situation will arise if a similar method is used to calculate the inertia-term correction.

It is convenient to have a shorthand notation for $1-J_0(\alpha_1 y i^{3/2})/J_0(\alpha i^{3/2})$. This will be written $F_0(\alpha_1 y)$. Its conjugate $F_0{}^*(\alpha_1 y)$ will also be needed. The full solution of (70) using this notation, is

$$w_1=\left(\frac{R^2}{\mu}\frac{A_1}{i^3\alpha^2}+\frac{1}{c_0}\frac{R^2}{\mu}\frac{A_0C_1}{i^3\alpha^2}\right)F_0(\alpha_1 y)$$
$$-\frac{\eta_1{}^*C_2A_1R^2}{c_0\mu 2i^3\alpha^2}\{F_0(\alpha_1 y)-F_0{}^*(\alpha_1 y)\}$$
$$+\frac{\eta_2C_1{}^*A_2R^2}{c_0\mu i^3\alpha^2}\{F_0(\alpha_1 y)-F_0(\alpha_2 y)\}. \quad . \quad . \quad . \quad . \quad . \qquad (72)$$

The expression for the average velocity can be written down at once from this, by replacing $F_0(\alpha_1 y)$ by $1-F_{10}$ wherever it appears, except in the main term, where it is to be replaced by $1+\eta_1 F_{10}$. If this is done. and the values of C_2 and $C_1{}^*$ inserted, we then form the expression

$$\overline{w}_1\,\mathrm{e}^{int}+\overline{w}_1{}^*\,\mathrm{e}^{-int}$$

which will be the first harmonic of the average velocity corresponding to the pressure gradient in its real form (60). We have

$$\begin{aligned}\overline{w}_1\,\mathrm{e}^{int}&+\overline{w}_1{}^*\,\mathrm{e}^{-int}\\&=-\frac{M_1}{n\rho}\,.\,M_{10}{}''\sin(nt+\phi_1+\epsilon_{10}{}'')-\frac{1}{c_0}\frac{M_0M_1}{n^2\rho^2}N_1M_{10}{}''M_{10}{}'\cos(nt+\psi_1)\\&\quad+\frac{1}{4c_0}\frac{M_2M_1}{n^2\rho^2}K_1N_1M_{20}{}''M_{10}{}'\{\cos(nt+\psi_2)-\cos(nt+\psi_3)\}\\&\quad-\frac{1}{c_0}\frac{M_2M_1}{n^2\rho^2}K_2N_1M_{10}{}''\{M_{10}{}'\cos(nt+\psi_4)-M_{20}{}'\cos(nt+\psi_5)\}\qquad .\;\;.\;\;.\quad(73)\end{aligned}$$

where we have written

$$\begin{aligned}\psi_1&=\phi_1-\delta_1+\epsilon_{10}{}''+\epsilon_{10}{}',\\\psi_2&=\phi_2-\phi_1-\delta_2-\gamma_1+\epsilon_{20}{}''+\epsilon_{10}{}',\\\psi_3&=\psi_2-2\epsilon_{10}{}',\\\psi_4&=\phi_2-\phi_1+\delta_1+\gamma_2-\epsilon_{10}{}''+\epsilon_{10}{}',\\\psi_5&=\psi_4+\epsilon_{20}{}'-\epsilon_{10}{}'.\end{aligned}$$

The main term in this expression is shown negative. This is correct. since the pressure gradient has been taken positive. The correction for the second harmonic can be obtained in the same way. It is convenient, when computing the corrections for a practical example, to divide throughout by the modulus of the main term, and this has been done for the correction to the second harmonic, which is shown below, with the common factor $1/npc_0$ omitted, and without the main term

$$\begin{aligned}M_0N_2M_{20}{}'&\cos(2nt+\psi_6)\\&+\frac{1}{4}M_1{}^2K_1N_1\frac{M_{10}{}''}{M_{20}{}''}\{M_{20}{}'\cos(2nt+\psi_7)-M_{10}{}'(\cos 2nt+\psi_8)\}\quad .\quad(74)\end{aligned}$$

where

$$\begin{aligned}\psi_6&=\phi_2-\delta_2+\epsilon_{20}{}''+\epsilon_{10}{}',\\\psi_7=\psi_8&=2\phi_1-\delta_1-\gamma_1+\epsilon_{10}{}''+\epsilon_{10}{}'.\end{aligned}$$

When $\sigma=\frac{1}{2}$ these expressions can be simplified slightly by substituting for K_1 and K_2, but this substitution is hardly worth making when computing. For numerical calculation, each correction is expressed as a set of coefficients of cos nt and sin nt, and these are summed and applied as proportional corrections to the components of the main term.

As an example of the magnitude of a typical set of corrections. a complete calculation has been done for an experiment by McDonald

(Womersley 1954) on the femoral artery of the dog. Fourier Analysis of his pressure gradient record gave

$$-\frac{\partial p}{\partial z}=0{\cdot}159+0{\cdot}774\cos(nt+0°\,39')+1{\cdot}317\cos(2nt-82°\,45')$$
$$-0{\cdot}743\cos(3nt+26°\,30')-0{\cdot}414\cos(4nt-16°\,39').$$

These coefficients are in mm of mercury per centimetre. The conversion constant to bring them to absolute units was included in the common factor $1/c_0 n\rho$.

It is not possible to make an accurate estimate of $1/c_0 n\rho$ until accurate measurements have been made of the pulse-velocity over short lengths of artery (see §6, below). The author was permitted to examine the film record of this experiment, and made a rough measurement of the maximum arterial expansion on the projected image. This gave a maximum $2\xi/R$ of about 6%. Since

$$\left|\frac{2\xi}{R}\right|=\frac{\overline{w}}{c_0}|x-iY|$$

and the maximum average velocity was 88 cm/sec, this gives

$$c_0=1310\ \text{cm/sec}$$

approximately. The pulse-velocity, estimated from records taken on other experiments, suggested a value of c_0 of about 850 cm/sec.

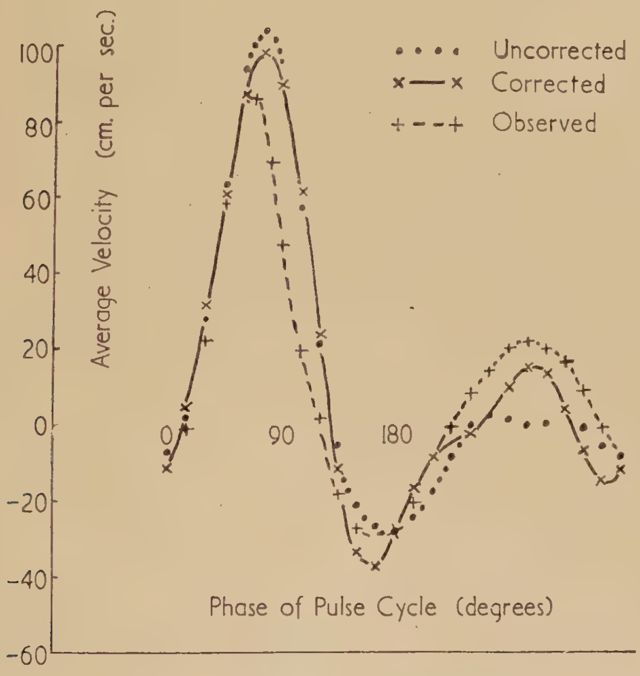
In the calculation of the corrections, $c_0=1000$ cm/sec was used, with $\sigma=\frac{1}{2}$. Table 4 gives a list of the coefficients for each harmonic, corrected and uncorrected. The effect on the steady component, which is not shown in the table, was to reduce it by about 2 cm/sec, rather more than 14%. This result is particularly important to the physiologist. It means that in a system such as this, in which the oscillatory terms are large, and the steady component small, Poiseuille's law cannot be used to calculate the steady flow from the pressure gradient—nor can the constant pressure gradient be inferred from the observed rate of flow or average velocity. Figure 5 shows the curves of average velocity over a complete pulse-cycle, plotted from the coefficients of table 4.

Table 4. Expansion-Correction
Fourier Coefficients for Calculation of Average Velocity

	Uncorrected		Corrected	
m	Coeff. of cos mnt	Coeff. of sin mnt	Coeff. of cos mnt	Coeff. of sin mnt
1	+19·08	+33·14	+19·07	+28·74
2	—31·78	+14·89	—35·13	+16·05
3	— 8·79	—10·58	—10·38	—10·99
4	— 0·44	— 5·86	+ 1·01	+ 3·29

The observed average velocity is also plotted, for comparison. It will be seen that the general shape of the curve is brought nearer to that of the observed curve by the correction. A rough calculation shows that to reduce the forward maximum on the calculated curve to the same as that for the observed curve it would be necessary to assume a wave-velocity of less than 500 cm/sec. This would not be consistent with the amount of arterial expansion observed.

Fig. 5



§ 6. CONCLUSION—THE NATURE OF THE APPROXIMATIONS

In the approximations so far made, the neglected terms are of order n^2R^2/c^2 or smaller, with the exception of the inertia terms in the equations of motion of the liquid. The largest of these might seem at first sight to be of order nR/c, but in fact they are larger than this, and their neglect is the major blemish on the theory outlined above. It is proposed to remedy this in Part II of this communication, and to make the approximation complete to order nR/c. Should it be necessary to take the approximation a stage further, and include terms of order n^2R^2/c^2, the term in $\partial^2 w/\partial z^2$ must be included, and the simple approximations $I_0(nRy/c)=1$, $I_1(nRy/c)=nRy/2c$ can no longer be used. The equation for the wave-velocity will then involve two non-dimensional parameters, α and nR/c. The larger values of nR/c which make this necessary are not likely to be important in the study of the circulation of the blood. In the larger animals increased values of R and lower values of c are offset by a lower pulse-rate. The range of values of α required to cover all conditions in the larger blood vessels in most animals is not very great, probably $1 \leqslant \alpha \leqslant 12$, though more accurate information is needed to check this general impression.

This does not mean that the study of the motion at higher frequencies is without interest. On the contrary, the question of the limit of stability of the motion of some importance, for the light it may throw on the incidence of turbulence in a straight pipe. No pipe is completely rigid, and something may be learnt about the incidence of turbulence by considering small perturbations of a steady pressure gradient, allowing for the elasticity of the walls.

To test the theory experimentally, the first requisite is an accurate measurement of pulse velocity in a short length of artery, say 3 or 4 cm, over which the diameter is reasonably constant. A method already exists with sufficient potential accuracy, but does not seem to have been used. The counter chronometer is now a standard piece of equipment, and such an instrument, working at one megacycle frequency, switched on and off by a pair of standard capacitance manometers, would be capable of measuring pulse-velocities to an accuracy better than one part in a thousand. Whether such accuracy could be achieved in practice is doubtful, owing to the rounded shape of the pulse-wave, making precise definition of its position difficult, but a manually adjustable d.c. gate, of the type used in some digital computers, could be used to define the crest of the pulse, and release a sharp-edged pulse from a trigger-circuit. Such measurements of pulse-velocity would be by far the most accurate ever made, and a detailed survey could be made of variations in pulse-velocity throughout the circulatory system.

The degree of damping could be measured in two ways, either by accurate recording of the form of the pulse-wave at two points a known distance apart, or by simultaneous recording of the pulse-wave and its derivative. If A is a complex Fourier Coefficient of the pressure and A' that of its space-derivative, then $A'=-inA/c$, and if c is complex, i.e. if $c_0/c=X-iY$, then $A'=A(X-iY)\,.\,in/c_0$, so that the phase-shift between the two will depend on the amount of damping. For measurements of this kind to be successful, the sensitivity of pressure-recording instruments must be raised. At present they are designed to measure pressures of the order of 100 mm of mercury, with a stable zero. The limits of the space-derivative measurement are of the order of ± 3 mm. If the amount of damping predicted by the theory proves to be greater than that observed, the simplest modification to take account of it would be to postulate different properties for the artery longitudinally and radially. To reduce damping, the artery would have to be freer longitudinally (so that its interaction with the moving mass of the fluid would be less) and stiffer laterally, a combination which is not compatible with homogeneous elastic properties. Such an elaboration is not justified until experimental proof is forthcoming that the degree of damping predicted is, in fact, too large, and the viscosity of blood is known with greater accuracy than at present, for the accuracy of the estimate of α depends upon this.

It will be shown in Part II that the corrections for the effect of the inertia terms, although appreciable, do not destroy the general picture given by the simple linear theory outlined above.

ACKNOWLEDGMENTS

The author acknowledges with thanks the help given by the Mathematics Division, National Physical Laboratory, in preparing the tables on which fig. 1 is based ; by Dr. J. F. Hale, Physiology Department, The Medical College of St. Bartholomew's Hospital, in computing tables 1–4, and by Miss E. Ullmann, of the same department, in preparation of diagrams.

To Professor K. J. Franklin, who welcomed and encouraged my continued interest in this work, and to Dr. D. A. McDonald, who originally aroused that interest and has given so freely of his time in advice and discussion, my warmest personal thanks are given.

The completion of this work was made possible by a Research Grant from the Medical Research Council, whose financial support is gratefully acknowledged.

REFERENCES

HELPS, E. P. W., and McDONALD, D. A., 1954, *J. Physiology*, **154**, 30P.
LAMB, H., 1898, *Mem. Manchester Lit. and Phil. Soc.*, **42,** No. 9.
LAMBOSSY, P., 1952, *Helvetica Physica Acta*, **25**, 371.
McLACHLAN, N. W., 1934, *Bessel Functions for Engineers* (Oxford).
MORGAN, G. W., and KIELY, J. P., 1954, *Journal of the Acoustical Society of America*, **26**, 323.
SHIPLEY, R. E., GREGG, D. E., and SCHROEDER, E. F., 1943, *Amer. J. Physiology*, **138,** 718.
WOMERSLEY, J. R., 1954, *J. Physiology*, **154,** 31P ; 1955, *Ibid.* (in press).

XXV. *Optical Constants of Metals in the Infra-red—Principles of Measurement*

By J. R. Beattie and G. K. T. Conn
Department of Physics, The University, Sheffield *

[Received November 2, 1954]

SUMMARY

In measuring the optical constants of an absorbing medium two pieces of information are required and may be obtained by examining the reflection of plane-polarized light at the principal angle of incidence. Since this angle is very close to 90° for metals in the infra-red, the value of measurements at smaller angles is examined. It is shown that there are serious disadvantages if reflecting coefficients only, such as $R_p{}^2$, $R_s{}^2$ or the ratio ρ^2, are measured. The ratio ρ of the reflecting coefficients of the p- and s-components and the relative phase difference Δ provide the most suitable parameters to measure at angles of incidence less than the principal angle. These are directly related to the modulus and the argument of the complex refractive index $n-ik$.

Calculation of optical constants from the measurements is always tedious. A chart has been prepared whereby the real and imaginary parts of the dielectric constant may be obtained from ρ and Δ with an accuracy of 1%; the preparation and use of these graphs is discussed. Illustrations are provided by quoting experimental results from $2\,\mu$ to $12\,\mu$ obtained with nickel, copper, silver, aluminium, mercury and antimony.

§ 1. INTRODUCTION

WHEN plane polarized radiation is reflected obliquely it becomes elliptically polarized. The relation between ϕ, the angle of incidence and χ the angle of refraction is given by Snell's law and the elementary conception of a real refractive index is extended for an absorbing medium so that an adequate description is given by a complex refractive index (Matossi 1949). Thus we have

$$\sin\phi = n\sin\chi = (n-ik)\sin\chi = n(1-i\kappa)\sin\chi.$$

The analytical justification of a complex refractive index is obvious yet its use requires care ; for example the surfaces of constant phase in the absorbing medium do not coincide with the surfaces of constant amplitude (Schuster 1904, Ditchburn 1952). Two pieces of information are required to describe the behaviour of an absorbing medium and therefore two measurements are necessary. The real and imaginary parts of the dielectric constant, n^2-k^2 and $2nk$, have a more direct significance than n and k particularly in dispersion.

---

* Communicated by Professor W. Sucksmith, F.R.S.

In recent years some attention has been given to measurement of the optical constants of absorbing materials and various techniques have been developed (Bock 1945, Simon 1951, Avery 1951, 1952, Robinson 1952, Robinson and Price 1953, Bor 1952, Archard, Clegg and Taylor 1952, Conn and Eaton 1954 a, b). This revival of interest is in part a new application of improved experimental methods and in part stimulated by the successes of the band theory of solids which provides a well established framework for interpretation. No attention has been given to the peculiar difficulties presented by the study of metals in the infra-red since the work of Försterling and Fréedericksz (1913). These difficulties have in the past proved such a stumbling block that, to the best of our knowledge, no reliable measurements with any metals of high conductivity have been reported at wavelengths longer than 5 μ. The purpose of the present paper is to indicate why experimental methods hitherto adopted fail when applied to metals in the infra-red and how the difficulties may be overcome. It emerges that the best parameters to measure are ρ and $\varDelta$. ρ is the ratio of the reflecting coefficients, R_p and R_s, of the p- and s-components of the reflected radiation and $\varDelta$ is the relative phase difference so that

$$\rho \exp(i\varDelta)=R_p/R_s=\cos\overline{\phi+\chi}/\cos\overline{\phi-\chi}. \quad . \quad . \quad . \quad . \quad (1)$$

The conventions adopted in defining R_p, R_s and $\varDelta$ are discussed briefly in an appendix since the accepted usage in optics is not self-consistent.

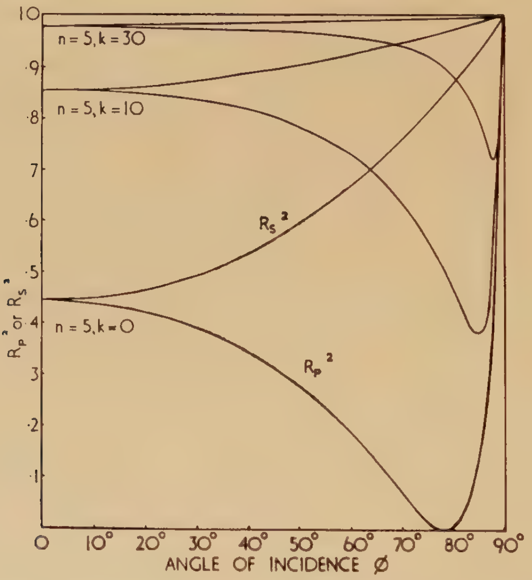
A laborious calculation must be made before n^2-k^2 and $2nk$ are obtained from measurements of ρ and $\varDelta$ at each wavelength. A great deal of time is saved if these operations are carried out graphically. This also ensures that the significance of changes in the experimental parameters is immediately grasped and the experiments can be pursued more intelligently. The most suitable graphical display is considered and extensively illustrated by practical examples.

§ 2. MEASUREMENT OF THE OPTICAL CONSTANTS OF METALS IN THE INFRA-RED

The optical constants of metals are best determined by studying reflected radiation. Examination of transmitted light is restricted by strong absorption of the conduction electrons to very thin specimens and the question arises to what extent the properties of such specimens accord with those of bulk material. In studies of reflection the relevant thickness is directly related to the ‘ penetration depth ’ and care must be given to the preparation of representative surfaces.

The two pieces of information required must be obtained from measurements which depend on different functions of n and k. If both measurements are associated with essentially the same function of n and k the experiments cannot be successful. R_p^2, or $R_pR_p^*$, and R_s^2, or $R_sR_s^*$, are plotted against ϕ in fig. 1, from which it appears that measurement in the neighbourhood of the principal angle $\bar{\phi}$ offers great promise and much attention has been given to such measurements. If metals are to be examined in the infra-red however, $\bar{\phi}$ usually exceeds 85° (Försterling and Fréedericksz 1913). For this spectral region the methods reported by Simon (1951), Avery (1951, 1952) and Conn and Eaton (1954 a) together with the method proposed by Abelès (1952) are of value in investigating materials of lower absorption when $\bar{\phi}$ is smaller but become very difficult when dealing with typical metals. The second method of Conn and Eaton (1954 b) is relevant but laborious for routine measurements. A new technique must be sought and the first step is to assess the benefit of working near $\bar{\phi}$ and the loss if ϕ is less than $\bar{\phi}$.

Fig. 1



Intensities reflected by various media as a function of the angle of incidence.

Ratios of intensities such as R_p^2, R_s^2 and ρ^2 are readily measured in the infra-red ; we consider R_s^2 and ρ^2 since the former obviously yields essentially the data obtained at normal incidence and the latter provides the additional information in a simple experimental form. R_0, the reflecting coefficient at normal incidence, is given by

$$R_0^2=\frac{\overline{n-1}^2+k^2}{\overline{n+1}^2+k^2} \text{ so that } \frac{1+R_0^2}{1-R_0^2}=\frac{n^2+k^2+1}{2n} . \quad . \quad . \quad . \quad (2)$$

Since both n^2 and k^2 are very much larger than unity it follows that measurement of R_0 yields $(n^2+k^2/2n)$. Relevant values of the optical constants are listed in tables 1 and 2.

Using the Fresnel expression discussed in the appendix it is easy to show that

$$R_s \exp(i\delta_S) = -\frac{s-\cos\phi}{s+\cos\phi} \text{ and } \rho \exp(i\Delta) = \frac{s-\tan\phi\sin\phi}{s+\tan\phi\sin\phi}, \quad . \quad . \quad (3)$$

where $\qquad s=\sqrt{(n^2-\sin^2\phi)}$ or $s^2=n^2-k^2-\sin^2\phi-2ink$.

If R_s^*, ρ^* and s^* are the complex conjugates of R_s, ρ and s respectively, so that $R_sR_s^*$ or R_s^2 is the coefficient of reflected intensity, we then have

$$A=\frac{1+R_sR_s^*}{1-R_sR_s^*}=\frac{ss^*+\cos^2\phi}{(s+s^*)\cos\phi}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

$$B=\frac{1+\rho\rho^*}{1-\rho\rho^*}=\frac{ss^*+\tan^2\phi\sin^2\phi}{(s+s^*)\tan\phi\sin\phi}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Equations (4) and (5) are exact and express measured parameters A and B in terms of the optical constants. If $\sin^2\phi$ is negligible compared with n^2, we may write ss^* as n^2+k^2 and $s+s^*$ as $2n$. Under these circumstances $\cos^2\phi$ is negligible compared with n^2+k^2 and the above relations are equivalent to

$$2A\cos\phi \doteqdot (n^2+k^2)/n \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

and

$$\begin{aligned}2B\tan\phi\sin\phi &\doteqdot (n^2+k^2+\tan^2\phi\sin^2\phi)/n \\ &=2A\cos\phi+\tan^2\phi\sin^2\phi/n. \quad . \quad . \quad . \quad . \quad (7)\end{aligned}$$

Equation (6), when compared with (2) shows that measurement of the reflected intensity of the s-component at any angle of incidence provides the same information as measurement at normal incidence, a conclusion which indeed is obvious from fig. 1. The second form of eqn. (7) reveals that unless the term $\tan^2\phi\sin^2\phi/n$ is comparable with $2A\cos\phi$ measurement of B gives no additional information which is significant. Conversely, provided ϕ is large enough to ensure that these terms are comparable, measurement may be made with profit. At the principal angle these two terms are equal since, with the approximations made above,

$$\tan^2\bar{\phi}\sin^2\bar{\phi}=n^2+k^2 \text{ (Drude 1887).} \quad . \quad . \quad . \quad . \quad (8)$$

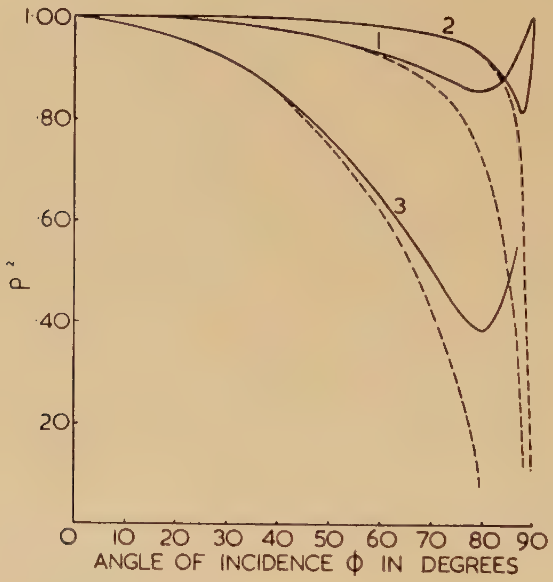
These conclusions, based on eqn. (7), are illustrated in fig. 2. $\rho\rho^*$ or ρ^2 is plotted against ϕ for the various cases listed in table 1. The continuous curves are obtained by calculating ρ^2 using both terms on the right hand side of eqn. (7). If the second term is ignored so that the right hand side is $2A\cos\phi$, the dotted curves are obtained ; these are therefore calculated from a knowledge of $(n^2+k^2)/n$ only and it is necessary to work at such an angle of incidence that these curves differ significantly. If, as in the method of Avery (1951, 1952), the required information is to be obtained by measurement of ρ^2 at two angles of incidence, one of these must be near the principal angle.

Table 1

	Wavelength	Material	n	k	Reference
1	$1{\cdot}03\,\mu$*	Cu (sputtered)	0·43	5·6	Försterling and Fréedericksz (1913)
2	$5{\cdot}10\,\mu$	Cu (sputtered)	2·98	28·4	„
3	$1{\cdot}00\,\mu$	Ni (bulk)	2·62	5·26	Ingersoll (1910)

* These values of n and k are conservative since values much larger than these are commonly obtained in the infra-red ; see table 2.

Fig. 2



ρ^2 or $R_p{}^2/R_s{}^2$ plotted against ϕ for the various cases listed in table 1. The numbers on the graphs refer to the order in table 1.

The measurements which lead to these conclusions are of $R_s{}^2$ and ρ^2 ; the difficulty arises because the moduli of both depend on the function $(n^2+k^2)/n$. To measure one of these is sufficient and ρ^2 is to be preferred for experimental reasons. The obvious associated parameter is the argument of ρ, that is Δ. It follows from eqn. (3), making the approximation that $\sin^2\phi$ is negligible compared with n^2, that

$$\tan\Delta=-\frac{2k\tan\phi\sin\phi}{n^2+k^2-\tan^2\phi\sin^2\phi}$$

hence

$$2\cot\Delta\tan\phi\sin\phi=-(n^2+k^2-\tan^2\phi\sin^2\phi)/k. \quad . \quad . \quad (9)$$

Equation (9) should be compared with eqn. (7). Since we are compelled to work at angles less than $\bar{\phi}$ the essential quality of the new equation is that $\cot\Delta$ is controlled by a new function $(n^2+k^2)/k$. If we ignore the second terms in (7) and (9), that is assume that $\tan^2\phi\sin^2\phi$ is much less than (n^2+k^2), an approximation which demonstrates the essential quality of these relations but which is too inaccurate to use in computing, we have that

$$n \sim (n^2+k^2)/(2B \tan\phi \sin\phi), \quad . \; . \; . \; . \; . \; . \; . \; . \quad (7\,a)$$

$$-k \sim (n^2+k^2)/(2 \cot \varDelta \tan\phi \sin\phi). \quad . \; . \; . \; . \; . \; . \; . \quad (9\,a)$$

Hence

$$|\,n\,| \sim 2 \tan\phi \sin\phi \Big/ \left(\frac{1}{B^2}+\tan^2 \varDelta\right) \text{ and } \arg n \sim B \tan \varDelta.$$

Measurement of ρ and $\varDelta$ offers a means of measuring the real and imaginary parts of the refractive index even if ϕ is significantly less than $\bar{\phi}$. Any two experimental parameters which give the required information must necessarily be related to the modulus and argument of the complex refractive index and therefore to ρ and $\varDelta$. The conclusion is therefore that ρ and $\varDelta$, or their equivalents, must be determined if measurements in the infra-red are to be successful. The experimental methods developed as a consequence of this are discussed in the following paper.

§ 3. COMPUTING THE OPTICAL CONSTANTS

To deduce numerical values of n and k or of n^2-k^2 and $2nk$ from measurements of ρ and $\varDelta$ requires laborious calculation. Two families of graphs have therefore been prepared whereby, knowing ρ and $\varDelta$, values of n^2-k^2 and $2nk$ may be obtained by inspection with an accuracy of the order of 1%. The use of these charts* is not of course limited to work in the infra-red. Moreover their value is not restricted to experiments in which ρ and $\varDelta$ are measured: in many cases transformation of the experimental variables to ρ and $\varDelta$ is simple (Conn and Eaton 1954 b).

From eqn. (3)

$$s/\tan\phi \sin\phi = (1+\rho \exp(i\varDelta))/(1-\rho \exp(i\varDelta)) \quad . \; . \quad (10)$$

Hence

$$\frac{s^2}{\tan^2\phi \sin^2\phi} = \frac{n^2-k^2-\sin^2\phi-2ink}{\tan^2\phi \sin^2\phi} = \left(\frac{1+\rho \exp(i\varDelta)}{1-\rho \exp(i\varDelta)}\right)^2 \equiv \tan^2 P \exp(2iQ). \quad . \; . \; . \quad (11)$$

By making use of complex conjugates it is readily shown that

$$\tan^2 P = \frac{1+\rho^2+2\rho \cos \varDelta}{1+\rho^2-2\rho \cos \varDelta} = \frac{1+y}{1-y} \text{ where } y = \frac{2\rho \cos \varDelta}{1+\rho^2}. \quad . \quad (12)$$

The phase angle is given by

$$\tan Q = \frac{2\rho \sin \varDelta}{1-\rho^2} = x \text{ so that } \sin 2Q = \frac{2x}{1+x^2} \text{ and } \cos 2Q = \frac{1-x^2}{1+x^2}. \quad (13)$$

From eqn. (11) we then have

$$X\,(n^2-k^2) \equiv \frac{n^2-k^2-\sin^2\phi}{\tan^2\phi \sin^2\phi} = \tan^2 P \,.\, \cos 2Q = \frac{1+y}{1-y}\cdot\frac{1-x^2}{1+x^2}. \quad (14)$$

$$\text{and} \quad Y\,(2nk) \equiv \frac{-2nk}{\tan^2\phi \sin^2\phi} = \tan^2 P \,.\, \sin 2Q = \frac{1+y}{1-y}\cdot\frac{2x}{1+x^2}. \quad (15)$$

---

*This is reminiscent of the well-known Smith charts used in studies of electrical impedance.

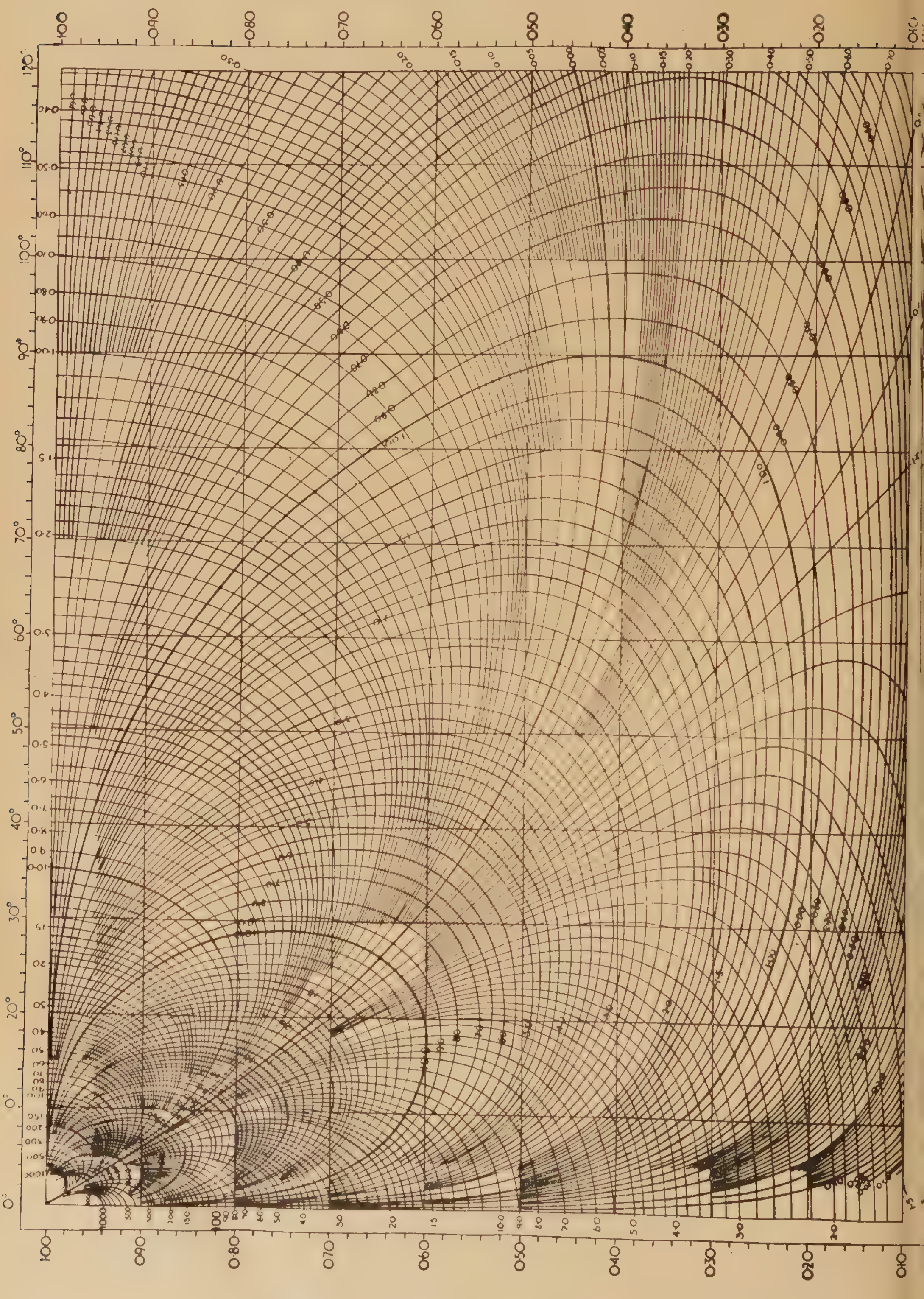

1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0°
10°
20°
30°
40°
50°
60°
70°
80°
90°
100°
110°
120°

Fig. 3

where y and x are the simple functions of ρ and Δ given respectively in (12) and (13). These equations, which are exact and which can be derived from the relations of Drude (1887), have been used to prepare families of contours or graphs.

The charts are shown on a greatly reduced scale in fig. 3. The abscissae are Δ, the relative phase, over the range 0° to 120°, which may be read on the scale used in practice with an accuracy of 0·2°. With the convention adopted (Appendix A), Δ is negative. The ordinates are the amplitude ratio, ρ, extending from 0·100 to 1·000; these may be read to 0·002. Since the practical values of n^2-k^2 and $2nk$ vary over a wide range, it is convenient to retain $\tan^2\phi\sin^2\phi$ in eqns. (14) and (15) as a scaling factor. When ϕ is 51° 48′, 73° 10′ or 84° 19′, $\tan^2\phi\sin^2\phi$ has the values 1, 10 or 100 respectively; it is therefore convenient to use one of these angles of incidence. If $\sin^2\phi$ is negligible, the presence of this scaling factor is the only influence of ϕ. One set of contours are the loci of constant values of

$$Y=-2nk/\tan^2\phi\,.\,\sin^2\phi\,;$$

these are the closed loops which terminate at the top left hand corner of the diagram with the particular value of Y indicated on each. The second set of contours are very roughly orthogonal to these and are loci of constant values of

$$X=(n^2-k^2-\sin^2\phi)/\tan^2\phi\,.\,\sin^2\phi.$$

Below the heavy zero line X is positive; above, it is negative. The measured values of ρ and Δ define a point; the values of X and Y are readily obtained by interpolation from the lattice of these curved contours. Over most of the chart, which measures 120 cm by 90 cm, X and Y may be obtained with an accuracy which varies between 1 % and 2 %.

In preparing these curves subsidiary graphs were first drawn on a large scale; these were of $\log_{10} X$ and $\log_{10} Y$ against ρ for various values of Δ and against Δ for various values of ρ. In those regions where $\log_{10} X$ and $\log_{10} Y$ become infinite, subsidiary graphs were drawn of X or Y against ρ and Δ. The accuracy of these must be somewhat greater than that available on the final chart and in all cases it was $\frac{1}{2}$%. In this way a large number of individual values of X and Y were plotted and the loci defined by particular values of X and Y were then drawn.

The practical accuracy for a range of materials and wavelengths is illustrated in table 2. The quoted values of the real and imaginary parts of the various dielectric constants have been measured in this laboratory and will receive fuller consideration in reports to be published shortly. In selecting samples for this table, an attempt has been made to illustrate the use of these charts over a wide range of ρ and Δ.

§ 4. General Comments

The merits of such a chart are not restricted to the speedy determination of $X(n^2-k^2)$ and $Y(2nk)$ following measurement of ρ and Δ. The accuracy of either n^2-k^2 or $2nk$ may be assessed by inspection. Thus when Δ is about 90° and ρ about 0·4, the precision with which ρ is measured fixes the accuracy of n^2-k^2 while Δ controls the accuracy of $2nk$. This may be contrasted with a case in which Δ is about 40° and ρ is about 0·3 when n^2-k^2 is primarily affected by Δ and ρ controls $2nk$. If particular precision is sought in either n^2-k^2 or $2nk$, a rough estimate of X and Y ensures that attention is focused on the relevant measurement. From the wider spacing of the contours it is clear that the highest accuracy is to be expected if Δ lies between 60° and 120°, that is in the region of the principal angle. Given the experimental errors in ρ and Δ, the errors in the values of n^2-k^2 and $2nk$ can be readily estimated.

Table 2

Specimen	λ in μ	ϕ	ρ	Δ	$2nk$		n^2-k^2		$\bar{\phi}$
					chart	calc.	chart	calc.	
Ni (bulk)	3·0	84°19′	0·699	80·2°	89·5	88·5	−103·5	−101·8	85°6′
	9·5	84°19′	0·870	36·8°	380	384	−780	−775	88°3′
Cu (bulk)	7·0	77°0′	0·973	12·3°	380	382	−1460	−1457	88°32′
Ag (bulk)	4·0	84°19′	0·948	49·5°	66·0	65·3	−463	−463·5	87°21′
Al (evap.)	2·0	84°19′	0·911	69·3°	41·5	41·5	−201·5	−203·6	86°3′
Hg (liquid)	6·0	83°30′	0·526	44·2°	300	297	−2	−3	86°41′
Sb (evap. metallic)	6·0	84°19′	0·219	111·3°	53·0	53·3	52·5	51·8	83°26′
		73°10′	0·506	18·2°	54·3	54·1	52·0	53·1	
		51°48′	0·806	5·2°	52·0	52·4	51·0	52·3	
	12·0	84°19′	0·496	101·2°	71·0	70·9	−17·6	−17·3	83°22′

The specimens referred to as ‘ evap.’ were prepared by evaporation *in vacuo*; the figures for antimony refer to the metallic phase. The figures listed in the column labelled ‘ calc.’ were the result of exact calculation.

For metals in the infra-red n^2-k^2 is almost invariably negative and $2nk$ is large. It follows from the former conclusion that ρ is greater than 0·42, and from the latter that Δ is small unless ϕ is large. Moreover if Δ is small, ρ is large ; the accuracy with which n^2-k^2 is obtained rests largely on Δ and of $2nk$ on ρ. The measured value of Δ may be increased and ρ reduced by studying multiple reflections.

§ 5. CONCLUSIONS

The initial stimulus of this investigation was to seek a satisfactory basis for measurement of the optical constants of metals in the infra-red region. Since ρ and Δ are essentially the modulus and argument of the complex refractive index these provide two different and distinctive pieces of information. This conclusion is of importance in any case where measurement at the principal angle of incidence is difficult or impossible.

The labour of preparing a computing chart such as that described is considerable ; a supply of working copies has been prepared and is available. With such a chart it is not only possible to convert experimental measurements rapidly to figures of the real and imaginary parts of the dielectric constant but appropriate steps can be taken to ensure that the measurements are made at suitable angles of incidence.

ACKNOWLEDGMENTS

In the course of the work we had many discussions with Dr. G. K. Eaton and it is a pleasure to acknowledge his assistance.

This work forms part of a programme for which we must acknowledge grants for apparatus from the Research Fund of the University of Sheffield and by the Royal Society.

APPENDIX A

The Phase Difference Δ and the Fresnel Coefficients

The expressions of Fresnel for the reflected amplitudes in the p- and s-directions, parallel and perpendicular respectively to the plane of incidence are usually written in the form (Jenkins and White 1951)

$$R_\mathrm{p}=\tan\overline{\phi-\chi}/\tan\overline{\phi+\chi}\ \text{ and }\ R_\mathrm{s}=-\sin\overline{\phi-\chi}/\sin\overline{\phi+\chi}. \qquad \text{(A.1)}$$

In this form there is a difference of sign interpreted as a change in phase of π for the s-component while there is no change in the phase of the p-component if $\overline{\phi+\chi}$ is less than 90°. At normal incidence however, the components are indistinguishable.

Derivations of R_s and R_p based on the elastic solid theory (Preston 1912) or on the concept of electrical impedance (Schelkunoff 1943) yield the same sign, contrary to eqn. (A.1). The simple experiments of Drude and Nernst (1892) show that light incident normally on a silvered surface is reflected with a change in phase of approximately π; the reflecting surface is a node. Both components must indeed undergo the same change in phase at normal incidence since Haidinger fringes are obtained as an unique set of bands. Moreover experiments at grazing incidence such as those of Wiener (1890) or those made with a Lloyd's mirror (Titchmarsh 1941) show that light of both components suffers a phase change of π in this case also.

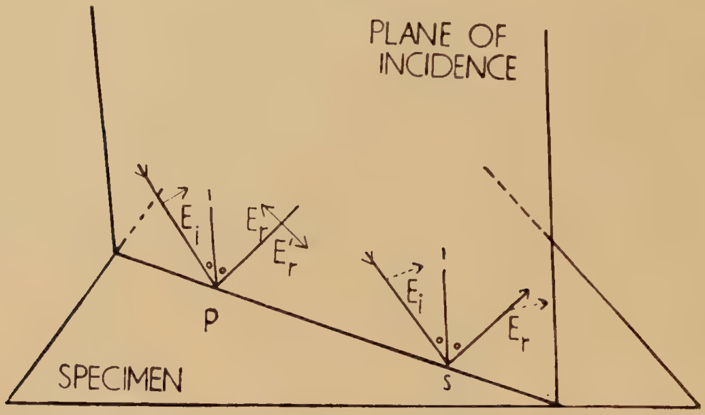
The discrepancy between these conclusions and eqn. (A.1) arises because of an inconsistency in definition. The positive vectors $\mathbf{E}_i$ and $\mathbf{E}_r$ illustrated in fig. 4 lead to the expressions (A.1). For the s-component both vectors are perpendicular to the plane of incidence and drawn in the same sense, a convention which conflicts with that adopted for the p-component when the angle of incidence is zero. This inconsistency is removed by adopting $\mathbf{E}_r{}^1$ in fig. 4 as the positive reflected vector associated with the p-component. It follows that

$$R_\mathrm{p}=-\tan\overline{\phi-\chi}/\tan\overline{\phi+\chi}\ \text{ and }\ R_\mathrm{s}=-\sin\overline{\phi-\chi}/\sin\overline{\phi+\chi}, \qquad \text{(A.2)}$$

and that

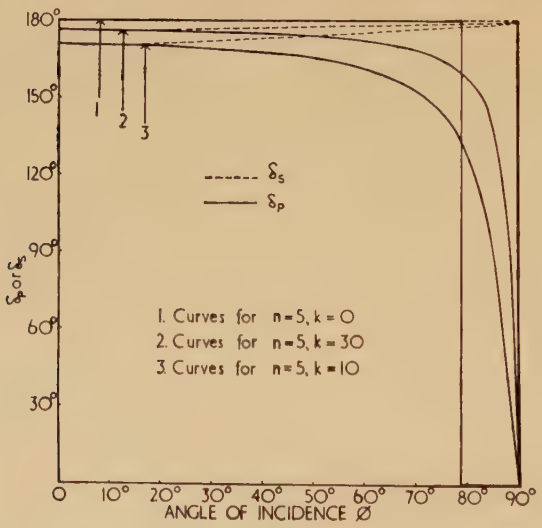
$$\rho\exp(i\Delta)=R_\mathrm{p}/R_\mathrm{s}=\cos\overline{\phi+\chi}/\cos\overline{\phi-\chi}. \qquad \text{(A.3)}$$

Fig. 4



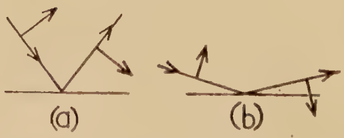
Illustrating the conventions adopted in defining the positive electric vector of a reflected ray.

Fig. 5



Phase changes at various angles of incidence. (*a*) for a dielectric, (*b*) for an absorbing material.

Fig. 6



The disposition of the positive reference vector when the angle of incidence is (*a*) small, (*b*) near grazing incidence.

With this convention both the p- and s-components suffer a phase change of π on reflection by a dielectric at normal incidence ; the phase difference Δ is zero. For an absorbing material it is easy to show that the reflecting coefficient at normal incidence and the associated phase changes are

$$R_0 = -\frac{n-ik-1}{n-ik+1} \text{ and } \tan \delta_{p,s} = -2k/(n^2+k^2+1). \quad . \quad . \quad (A.4)$$

Δ is zero as before and in the infra-red where n^2 and k^2 are both large compared with unity, $\delta_{p,s}$ is but slightly less than π.

Figure 5 shows the phase changes at various angles of incidence for a dielectric and for absorbing materials. It appears that while the s-component suffers a phase change of π at grazing incidence, the p-component does not. It is obvious, however, from fig. 6 that the positive sense of the reflected p-vector leads to destructive interference at the surface of the mirror ; figs. 6 (*a*) and 6 (*b*) show the positive reference vector at angles of incidence which are (*a*) small and (*b*) near 90°. The direction of reference moves through 90° as the angle of reflection changes from 0 to 90° so that the positive sense of the reflected vector is inverted when compared with that of the incident vector. This inversion in sense of the reference vector must be borne in mind.

APPENDIX B

The Use of Data from Measurements of Reflectivity

If measurements of reflected intensities are used to determine the real and imaginary parts of the dielectric constants, we have from the exact eqns. (4) and (5),

$$|s|^2 - 2ACr(s) + C^2 = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (B.1)$$

and

$$|s|^2 - 2BTr(s) + T^2 = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (B.2)$$

In these equations $r(s)$ is the real part of s. C is written for $\cos\phi$ and T for $\tan\phi \, . \, \sin\phi$. The solutions of these equations are

$$2nk = 2l\,(m-l^2)^{1/2} \text{ and } n^2-k^2 = 2l^2 - m + \sin^2\phi,$$

where

$$m = |s|^2 = TC(AT-BC)/(BT-AC) \text{ and } l = r(s) = \tfrac{1}{2}(T^2-C^2)/(BT-AC).$$

These expressions although new in form are equivalent to those given by Abelès (1952). They imply measurement of R_s^2 and ρ^2 at a particular value of ϕ.

The convenient experimental method of Avery (1951, 1952) requires a solution of $B2$ for different values of ϕ. This is possible only by graphical means or by assuming that $\sin^2\phi$ is the same for both values of ϕ (Hodgson 1954).

REFERENCES

Abelès, F., 1952, *Rev. d'Opt.*, **31,** 127.

Archard, J. F., Clegg, P. L., and Taylor, A. M., 1952, *Proc. Phys. Soc.* B, **65,** 758.

Avery, D. G., 1951, *Proc. Phys. Soc.* B, **64,** 1087 ; 1952, *Ibid*, **65,** 425.

Bock, R. O., 1945, *Phys. Rev.*, **68,** 210.

Bor, J., 1952, *Proc. Phys. Soc.* B, **65,** 753.

Conn, G. K. T., and Eaton, G. K., 1954 a, *J. Opt. Soc. Amer.*, **44,** 484 ; 1954 b, *Ibid.*, **44,** 546.

Ditchburn, R. W., 1952, *Light* (London : Blackie), p. 480.

Drude, P., 1887, *Ann. Phys. Lpz.*, **32,** 584.

Drude, P., and Nernst, W., 1892, *Ann. Phys. Lpz.*, **45,** 460. (See also Wood, R. W., 1924, *Phys. Optics*, 2nd Ed. (Macmillan), p. 176.

Försterling, V., and Fréedericksz, V., 1913, *Ann. Phys. Lpz.*, **40,** 201.

Hodgson, J. M., 1954, *Proc. Phys. Soc.* B, **67,** 269.

Ingersoll, L. R., 1910, *Astrophys. Jnl.*, **32,** 265.

Jenkins, F. A., and White, H. E., 1951, *Fundamentals of Optics*, 2nd Ed. (McGraw-Hill).

Matossi, F., 1949, *J. Opt. Soc. Amer.*, **39,** 506.

Preston, T., 1912, *Theory of Light*, 4th Ed. (Thrift, Macmillan), pp. 362, 364.

Robinson, T. S., 1952, *Proc. Phys. Soc.* B, **65,** 910.

Robinson, T. S., and Price, W. C., 1953, *Proc. Phys. Soc.* B, **66,** 969.

Schelkunoff, F. A., 1943, *Electromagnetic Waves* (Van Nostrand), p. 258.

Schuster, A., 1904, *Theory of Optics* (Arnold).

Simon, I., 1951, *J. Opt. Soc. Amer.*, **41,** 336.

Titchmarsh, P. T., 1941, *Proc. Phys. Soc.*, **53,** 391.

Wiener, O., 1890, *Ann. Phys. Lpz.*, **40,** 203. (See also Wood's *Optics*, p. 175.)

XXVI. *Optical Constants of Metals in the Infra-red—Experimental Methods*

By J. R. Beattie
Department of Physics, The University, Sheffield*

[Received November 2, 1954]

Summary

For metals the principal angle of incidence in the infra-red is near 90° and it has been shown that the real and imaginary parts of the dielectric constant are best obtained by measuring ρ and Δ ; ρ is the ratio of the reflectivities of the p- and s-components and Δ the relative phase difference. ρ and Δ may be obtained without using birefringent materials by measuring two ratios of intensities, a technique which provides a cross-check. The experimental methods and arrangements developed to measure a number of metals in the spectral range $2\,\mu$ to $12\,\mu$ are described ; this range can be extended where necessary. Use of multiple reflections is discussed and errors are examined. Measurements are reported of evaporated aluminium mirrors since these are frequently used.

§ 1. Introduction

In the sixty years which have elapsed since the definitive work of Drude on the measurement of optical constants numerous authors have reported results of experiments on metals but very few of these refer to the infra-red region. The measurements of Försterling and Fréedericksz (1913) extend to $5\,\mu$, those of Ingersoll (1910), Bock (1945) and Bueche (1948) to $2{\cdot}5\,\mu$. Data on the reflectivity at longer wavelengths are available, largely the work of Hagen and Rubens (1904), but such measurements of a single parameter are of limited interpretative value. The figures of Försterling and Fréedericksz have necessarily received most general acceptance but as Wilson (1936 and 1953) has observed there are serious discrepancies with theory.

The success of the modern theory of metals and developments in experimental technique both emphasize the need for a study of the optical constants of metals in the infra-red. The problems peculiar to such measurements are examined in a paper (Beattie and Conn 1955) which shows why previous experiments have been confined to relatively short wavelengths and the manner in which these difficulties may be overcome. The measurements required of radiation reflected by a specimen are ρ, the ratio of the reflectivities R_p and R_s of the p- and s-components, and Δ the relative phase difference. The technique is thus an adaptation of one of Drude's methods to suit present limitations in the infra-red. Since suitable birefringent materials are not available optical elements which give plane polarized light must be used.

---

* Communicated by Professor W. Sucksmith, F.R.S.

In the present paper the experimental methods which have been developed are discussed and sources of error are examined. Practical experience has been obtained in studies of silver, copper, nickel, antimony and mercury over the range of $2\,\mu$ to $12\,\mu$ and there is every prospect of extending such measurements, where necessary, to longer wavelengths. Detailed reports on these metals are now in preparation. Practical illustrations are given in the light of this experience and by reporting measurements of the optical constants of aluminium mirrors prepared by evaporation *in vacuo* since these are widely used.

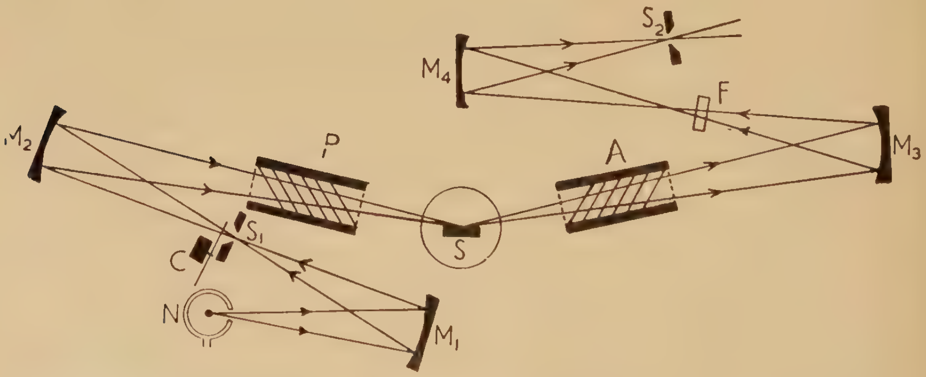
§ 2. EXPERIMENTAL METHOD

Plane polarized radiation is obliquely reflected by a specimen and passes through an analyser to the monochromator and detector. The intensity transmitted by the system ASP of fig. 1, $I(\psi_P, \psi_A)$ for azimuths ψ_P and ψ_A of the polarizer and analyser respectively is of the form (Conn and Eaton 1954 a),

$$I(\psi_P, \psi_A)=I_0(\sin^2 \psi_P \sin^2 \psi_A+\rho^2 \cos^2 \psi_P \cos^2 \psi_A+\tfrac{1}{2}\rho \sin 2\psi_P \sin 2\psi_A \cos \Delta). \quad . \quad . \quad . \quad (1)$$

For metals of high conductivity R_s differs little from unity and is conveniently incorporated in $I_0=\frac{1}{2}E^2R^2{}_s$; E is the amplitude of the incident light.

Fig. 1



Optical path from the source N to the entrace slit of the monochromator S_2. A, S and P represent respectively an analyser, a specimen and a polarizer (see text).

Measurements of the transmitted signal are required at three different azimuths at least to obtain ρ and Δ. At a fixed ψ_P the intensities I_1 and I_2 are measured when the analyser is perpendicular and parallel to the plane of incidence so that

$$I_1\equiv I(\psi_P, \pi/2)=I_0 \sin^2 \psi_P\,; \quad I_2\equiv I(\psi_P, 0)=I_0\rho^2 \cos^2 \psi_P.$$

Hence

$$\rho=\tan \psi_P \sqrt{(I_2/I_1)}. \quad . \quad . \quad . \quad . \quad (2)$$

ψ_P is chosen to ensure accurate measurement of ρ. To measure Δ the third term on the right hand side of (1) should clearly be as large as possible; the azimuths should therefore be $\pm\pi/4$ and thus

$$I_3 \equiv I(\pm\pi/4, \pm\pi/4) = I_0(1+\rho^2+2\rho\cos\Delta)/4. \quad . \quad . \quad . \quad . \quad (3)$$

Then

$$\cos\Delta = \{2I_3/I_1 - (1+\rho^2)\}/2\rho. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Also

$$I_4 \equiv I(\pm\pi/4, \mp\pi/4) = I_0(1+\rho^2-2\rho\cos\Delta)/4 \quad . \quad . \quad . \quad . \quad (5)$$

and

$$\cos\Delta = \{(1+\rho^2) - 2I_4/I_1\}/2\rho. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

In principle either (4) or (6) may be used to measure Δ but neither is satisfactory. Since both ρ and Δ are required to provide two pieces of information such as the real and imaginary parts of the complex refractive index, it is desirable that the measurements are independent which (4) and (6) are not. If (4) and (6) are used together it follows that

$$\cos\Delta = \tfrac{1}{2}\left(\rho + \frac{1}{\rho}\right)\{(I_3 - I_4)/(I_3 + I_4)\}. \quad . \quad . \quad . \quad . \quad (7)$$

With the values of ρ usually encountered $\frac{1}{2}[\rho + (1/\rho)]$ does not depart seriously from unity; for example if ρ is 0·8 the value is 1·025. In general

$$\text{if} \qquad 0{\cdot}40 \leqslant \rho \leqslant 1 \quad \text{then} \quad 1{\cdot}45 \geqslant \tfrac{1}{2}\left(\rho + \frac{1}{\rho}\right) \geqslant 1.$$

Equation (7) offers a means of measuring Δ largely independent of ρ, a feature which is particularly relevant to the small values of Δ found at long wavelengths when ρ tends to unity.

Equation (1) is symmetrical in ψ_A and ψ_P so that the polarizer and analyser are interchangeable. This is a matter of some experimental importance since a prism monochromator shows such pronounced polarization that it has been used to estimate dichroism (Hyde 1948). With the simple prism instrument used in these experiments the ratio of the transmitted intensities at 2 μ for radiation polarized in the vertical and horizontal planes is about 0·70. The difficulty is eliminated if the azimuth of the polar between specimen and monochromator is fixed, that is, it fulfils the role of polarizer (ψ_P). The first polar is then treated as the analyser (ψ_A).

§ 3. Azimuths of the Polarizer and Analyser

ψ_P and ψ_A are measured with respect to the principal plane of the specimen and the origins of the two azimuth scales can be located geometrically with an accuracy of about 1° of arc. Greater precision is essential if systematic errors are to be avoided. The zero errors $\delta\psi_P$ and $\delta\psi_A$ of these scales can be measured by examining any convenient specimen at such an angle of incidence that Δ is approximately $\pi/2$ and at such a wavelength that the signal is large. Nickel has been used at an angle of incidence of about 84° and a wavelength of 2·2 μ; these values are suitable for hand-polished, electro-polished or buffed nickel.

In the first experiment to measure $\delta\psi_A$ the polarizer is set at a nominal azimuth of $\pi/2$ and intensities are compared for two settings $\pm\psi_A'$ of the analyser. Inequality of the two signals reveals a systematic error in the setting of the analyser.

If $$\psi_P=\pi/2-\delta\psi_P \text{ and } \psi_A=\psi_A'-\delta\psi_A,$$
eqn. (1) gives to a first order approximation,
$$I(\pi/2-\delta\psi_P, \psi_A'-\delta\psi_A)\doteqdot I_0 \sin^2\psi_A' \{1-2\cot\psi_A'(\delta\psi_A-\rho\cos\Delta\,.\,\delta\psi_P)\}$$
and
$$I(\pi/2-\delta\psi_P,-\psi_A'-\delta\psi_A)\doteqdot I_0 \sin^2\psi_A' \{1+2\cot\psi_A'(\delta\psi_A-\rho\cos\Delta\,.\,\delta\psi_P)\}.$$
$\delta\psi_P$ disappears if Δ is $\pi/2$ so that
$$\delta\psi_A=(1-\alpha)/4\cot\psi_A'$$
where α is the ratio of the measured intensities. If ψ_A' is between 30° and 60°, $\delta\psi_A$ may be measured to $\pm 0{\cdot}1°$.

After applying this correction so that $\delta\psi_A$ is zero, $\delta\psi_P$ may be measured in a number of ways. The most convenient is probably to examine discrepancies in measurements of ρ^2. Following eqn. (2) we define
$$\beta\equiv I(\pi/4-\delta\psi_P, 0)/I(\pi/4-\delta\psi_P, \pi/2)$$
and
$$\gamma\equiv I(-\pi/4-\delta\psi_P, 0)/I(-\pi/4-\delta\psi_P, \pi/2).$$
Ignoring second and higher order infinitesimals it follows from eqn. (1) that
$$\beta\doteqdot\rho^2(1+4\delta\psi_P) \text{ and } \gamma\doteqdot\rho^2(1-4\delta\psi_P)$$
hence
$$\delta\psi_P=(\beta/\gamma-1)/8.$$
In practice $\delta\psi_P$ may also be measured to $\pm 0{\cdot}1°$. It is a merit of this technique that errors in azimuth may be measured using specimens under examination. This is indicated above by quoting figures for nickel.

§ 4. THE REFLECTANCE RATIO

In examining the use of eqn. (2) to measure ρ, errors in azimuth will be neglected since the most important source of random error is fluctuation of the signal δI.
$$\delta\rho/\rho=\tfrac{1}{2}\delta I(1/I_1+1/I_2)=\tfrac{1}{2}\delta I(\operatorname{cosec}^2\psi_P+\sec^2\psi_P/\rho^2)/I_0.$$
The smallest value of $\delta\rho/\rho$ occurs when $\sqrt{\rho}=\tan\psi_P$; then
$$\delta\rho/\rho=\delta I\left(1+\frac{1}{\rho}\right)^2\Big/2\,.\,I_0.$$
This optimum condition requires a change of ψ_P with each specimen and it is experimentally more convenient to set $\psi_P=\pi/4$ when
$$\delta\rho/\rho=\delta I\left(1+\frac{1}{\rho^2}\right)\Big/I_0. \qquad \ldots\ldots \quad (8)$$
Taking as a practical figure 1/500 for $\delta I/I_0$, eqn. (8) indicates that $\delta\rho/\rho$ is about 0·4% when ρ is of the order of unity and 1·5% when ρ is as low as 0·4; if the optimum condition is obeyed the corresponding figures are 0·4% and 1·2%. In the infra-red ρ is relatively large (Beattie and Conn 1955) and tends towards unity as the wavelength increases.

§ 5. THE RELATIVE PHASE

The influence of noise may be assessed by writing unity in eqn. (7) for $\frac{1}{2}[\rho+(1/\rho)]$ so that

$$\cos \varDelta \div (1-I_4/I_3)/(1+I_4/I_3) \text{ and } \delta\varDelta = \delta(I_4/I_3)\sqrt{(I_4/I_3)}/(1+I_4/I_3).$$

Since errors in I_4 and I_3 must be added it follows that

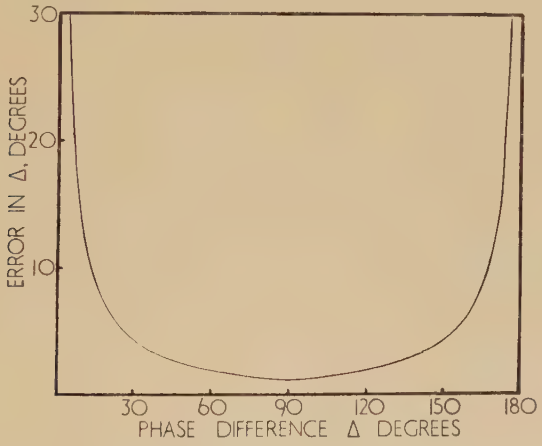
$$\delta\varDelta = \frac{\delta I}{\sqrt{(I_3 \,.\, I_4)}} = \frac{\delta I}{I_3} \Big/ \sqrt{\left(\frac{I_4}{I_3}\right)} = \frac{\delta I}{I_4} \Big/ \sqrt{\left(\frac{I_3}{I_4}\right)}. \quad . \quad . \quad . \quad (9)$$

The second form of eqn. (9) is suitable for $\varDelta \leqslant \pi/2$ when we write

$$\delta I/I_3 = 1/500$$

as in § 4; the third form is used if $\varDelta \geqslant \pi/2$ when $\delta I/I_4$ is set equal to 1/500. In effect the assumption is made that either I_3 or I_4, whichever is relevant, gives a full scale deflection.

Fig. 2



Error in the relative phase $\varDelta$ plotted against $\varDelta$.

The error in $\varDelta$ is shown as a function of $\varDelta$ in fig. 2 where $\delta\varDelta$ is given by

$$\delta\varDelta = 1/(500\sqrt{a}) \text{ with } a = (1-\cos \varDelta)/(1+\cos \varDelta).$$

The need to work at reasonably large angles of incidence is obvious since the accuracy deteriorates when $\varDelta$ is less than about 15°. This is inherent since it is $\cos \varDelta$ which appears in eqn. (1), but multiple reflections may be used to overcome the limitation. This technique is discussed below. In practice 85° is the maximum angle of incidence which can be used and as the wavelength increases $\varDelta$, of necessity therefore, tends to zero.

§ 6. GENERAL COMMENTS

Four measurements of the signal are thus made at each wavelength, ψ_{P} being $\pi/4$ in all cases. Two of these, $\psi_{\mathrm{A}}=0$ and $\pi/2$, give ρ using eqn. (2); two, $\psi_{\mathrm{A}}=\pm\pi/4$, give $\varDelta$ using eqn. (7). ρ and $\varDelta$ are parameters of an idealized ellipse, that is of the elliptically polarized radiation reflected by the specimen. Measurement in practice gives values averaged over a range of wavelengths and over a range of angles of incidence. If the radiation is brought to a focus on the specimen the surface area may be reduced without loss of flux. Since only three of the four measured intensities are necessary it is possible to check that the range of wavelengths and of incidence angles is not excessive. From eqns. (2), (3) and (5) it is easy to show that

$$I_1+I_2=I_3+I_4.$$

It emerges from experience that this equation should be obeyed to within 1 or 2%.

The symmetry of eqn. (1) is used from time to time to check the zeros of the azimuth scales. These may alter by accident or because of sagging or cracking of the selenium films which form the polars. The relation

$$I(\psi_{\mathrm{P}}, \psi_{\mathrm{A}})=I(-\psi_{\mathrm{P}}, -\psi_{\mathrm{A}}) \text{ or } I(+)=I(-)$$

should be obeyed. If it is not obeyed,

$$\begin{aligned} I(\psi'_{\mathrm{P}}, \psi'_{\mathrm{A}}) &= I(\psi_{\mathrm{P}}+\delta\psi_{\mathrm{P}}, \psi_{\mathrm{A}}+\delta\psi_{\mathrm{A}}) \\ &= I(\psi_{\mathrm{P}}, \psi_{\mathrm{A}})+(\delta I/\delta\psi_{\mathrm{P}})\delta\psi_{\mathrm{P}}+(\delta I/\delta\psi_{\mathrm{A}})\delta\psi_{\mathrm{A}} \qquad (10) \end{aligned}$$

and

$$\begin{aligned} I(-\psi'_{\mathrm{P}}, -\psi'_{\mathrm{A}}) &= I(-\psi_{\mathrm{P}}+\delta\psi_{\mathrm{P}}, -\psi_{\mathrm{A}}+\delta\psi_{\mathrm{A}}) \\ &= I(\psi_{\mathrm{P}}, \psi_{\mathrm{A}})-(\delta I/\delta\psi_{\mathrm{P}})\delta\psi_{\mathrm{P}}-(\delta I/\delta\psi_{\mathrm{A}})\delta\psi_{\mathrm{A}}. \qquad (11) \end{aligned}$$

By averaging the signals $I(+)$ and $I(-)$ the influence of these errors in azimuth is eliminated at the expense of doubling the number of readings. This means of dealing with such errors, not that described in § 3, was used in measurements of liquid mercury.

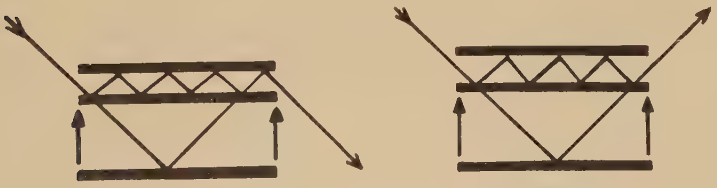
§ 7. EXPERIMENTAL DETAILS

The source of radiation, a Nernst glower N in fig. 1, is inside a water jacket to reduce thermal fluctuations of intensity. Light emitted obliquely from the edges of a filament is partially polarized (Laüe and Martens 1907, Czerny 1924); a measurement of the ratio of intensities of such light, polarized in vertical and horizontal planes, gave about 0·80. The beam is restricted to radiation from the centre of the source by focusing a sharp image on the variable slit S_1 and to verify that the light is then substantially free from polarization, the width of S_1 should not influence ρ^2. Alternatively the specimen is removed and the apparent value of ρ^2 for ‘ free space ’ is measured ; this should not differ significantly from unity for all wavelengths. The second alternative is more thorough but the polars and the mirrors M_2 and M_3 must be rearranged.

The efficiency of the polarizing elements must be high. A percentage polarization of at least 99·5% is obtained with each polar using six films of selenium inclined to the beam at the Brewster angle (Elliot, Ambrose and Temple 1948). The behaviour of such polarizing devices has been examined by Conn and Eaton (1954 b) who showed the importance of eliminating multiple reflections between films. In the present arrangement the effect of these reflections is removed by carefully centring the image on the entrance slit of the monochromator, S_2. Since the azimuth of the analyser is changed from one reading to the next the beam must be clearly defined by a stop to prevent fouling by the metal frame of the polar. This is the purpose of a circular diaphragm on the mirror M_1, imaged by M_2 on the aperture of the analyser. The convergence of light on the specimen, which is placed at the focus conjugate to S_1, is controlled by a rectangular diaphragm at M_2.

A turntable, free to move about a vertical axis and graduated in degrees and minutes, supports a specimen; the plane of the front surface is defined by three ball contacts. This simple arrangement is adequate to examine light reflected once. Without increasing the angle of incidence, the accuracy may be improved at longer wavelengths, when ρ tends to unity and Δ becomes small, by studying m reflections between a pair of specimens. ρ^m and $m\Delta$ are then measured. Figure 3 (*a*) illustrates an arrangement whereby the number of reflections may be changed without altering the optical path (see Jamin 1847), and since two reflections are excessive for accurate measurement of the relatively low values of ρ in the neighbourhood of $2\,\mu$, fig. 3 (*b*) shows the case which permits study of single reflections. In both cases the mirrors are mounted on separate carriages both of which may move on a parallel slide clamped to the turntable; the mechanical quality of the slide and of these carriages is clearly important. The radiation is brought to a focus between the mirrors, the half-angle of the converging cone being about $\frac{1}{2}°$ or an aperture ratio of about 1/50. Serious errors arise if ϕ is greater than 85° since the angle of incidence affects the calculated values of the real and imaginary parts of the dielectric constant by the presence of the factor $\tan^2\phi\,\sin^2\phi$ (Beattie and Conn 1955). At greater angles parallel illumination would be essential but the width of the source and of the spectrometer slit vitiate attempts to define ϕ more accurately. In practice the angle of incidence is chosen and by altering the separation of the plates the number of reflections is increased to compensate for the diminution

Fig. 3



(*a*) (*b*)

Arrangement for the study of multiple reflections.

of Δ and $(1-\rho)$ at long wavelengths. Multiple reflections have been of value for measurements of silver, copper, nickel and aluminium beyond 6 μ; for wavelengths less than about 12 μ and all save metals of the highest conductivity single reflection and a single specimen one or two inches long are adequate.

Since there is no cogent reason to anticipate that the optical constants of metals change rapidly with wavelength in the infra-red, a monochromator of modest resolution has been used hitherto. As in the arrangement of McAlister (1941) the radiation passes twice through a 60° rock-salt prism of base 3 cm. Over the range 2 μ to 14 μ the resolution is of the order of 0·1 μ but a slit width as large as 0·25 μ or even 0·5 μ at long wavelengths has been used because of the intensity lost by defining apertures and polarizers before the light enters the monochromator. The instrument has been calibrated and checked from time to time by using the well-known atmospheric absorption bands and cells containing benzene and chloroform. The effects of radiation scattered within the monochromator can be serious particularly on measurements of I_4. At long wavelengths I_4 is small and it follows that a large proportion of the signal at 10 μ, say, may arise from scattered radiation of shorter wavelengths. Although attention has been given to this in the design of the spectrometer (Williams 1948), precautions must be taken against the possibility that about 1% of the signal at wavelengths greater than 9 μ arises from radiation of short wavelength. For measurements at wavelengths longer than 6 μ a selenium powder filter is used (Pfund 1933); this filter, indicated by F in fig. 1, transmits 4% at 2 μ and 90% at 6 μ. At wavelengths greater than 9 μ a mica sheet is substituted for the opaque chopper blade commonly used. As the result of these precautions it is believed that the data obtained are substantially free from the effects of light scattered in the monochromator.

The beam is interrupted at a frequency of 2·67 cycles per second by a chopper C in fig. 1. The e.m.f. of the evacuated Schwarz thermocouple is magnified by a resonant galvanometer amplifier tuned to this frequency by adjusting the feed-back; this is followed by an electronic amplifier (see Beattie and Conn 1954). The output is presented, rectified and smoothed, on a linear scale of 500 divisions.

The longest wavelength at which measurements have been made to date is 13 μ. This limit is set by intensity of the source and diminishing values of Δ. I_4 decreases and accuracy deteriorates beyond 10 μ. At shorter wavelengths changes of 1% in (n^2-k^2) and $2nk$ can usually be determined. Since ϕ is fixed, cross-checks ensure that only minor errors affect ρ and Δ but absolute accuracy is affected by measurement of the mean value of ϕ; an error of 0·1° for instance at 85° causes a systematic error of 4% in (n^2-k^2) and $2nk$. Though there may in consequence be some uncertainty in the absolute values, the variation of the dielectric constant with wavelength is not affected. This source of error may be reduced by decreasing ϕ which is possible using multiple reflections.

§ 8. Optical Constants of Aluminium Mirrors

Mirrors prepared by evaporation *in vacuo* of aluminium condensed on glass are very widely used, particularly in the infra-red. It is therefore important to record the optical constants of such mirrors which are frequently used as a reference in measurements of reflectivity. Moreover in experiments such as these under discussion appreciable polarization may be introduced by mirrors; for instance M_1 and M_2 in fig. 1 are not used on their optic axes and may therefore contribute significantly to apparent polarization from the source. Measurements of such specimens reveal the orders of magnitude encountered in examining metals of high conductivity and give practical point to many of the comments made above.

Evaporated aluminium was condensed at 5 to 10 Å.U. per second on two identical glass plates at a distance of 20 cm from the source. The residual gas pressure within the chamber during evaporation was of the order of 10^{-5} mm Hg; a liquid nitrogen trap was used between this chamber and the oil pumps. The thickness of the film, as measured subsequently by interferometry, was 1550 ± 50 Å.U. The optical constants of these surfaces were measured from $2\,\mu$ to $12\,\mu$ within twenty-four hours of their formation. The angle of incidence was 84·2° and up to five reflections were used in covering this range. The results are listed in the table.

Optical Constants of Mirrors of Evaporated Aluminium

λ in μ	ρ	$\Delta°$	n^2-k^2	$2nk$	n	k	$R_0{}^2$
2·0	0·885	60·8	−268	76	2·30	16·5	0·968
2·5	0·885	50·7	−401	130	3·22	20·3	0·970
3·0	0·883	43·0	−565	213	4·41	24·2	0·971
4·0	0·894	34·6	−881	362	5·97	30·3	0·975
5·0	0·899	28·5	−1285	604	8·19	36·8	0·978
6·0	0·898	24·6	−1670	930	11·00	42·4	0·977
7·0	0·900	20·8	−2185	1440	14·63	49·0	0·978
8·0	0·907	18·5	−2730	1870	17·00	55·0	0·980
9·0	0·909	16·2	−3305	2590	21·1	61·3	0·980
10·0	0·910	14·5	−3890	3405	25·4	67·3	0·980
11·0	0·910	13·6	−4270	4025	28·3	71·1	0·981
12·0	0·910	12·3	−4700	5130	33·6	76·4	0·981

The absolute accuracy to be attached to values of n^2-k^2 and $2nk$, of the order of 4%, is determined by the error in ϕ and therefore in $\tan^2\phi\sin^2\phi$. Since k is much larger than n it follows that the error in k is of the order of 2% and, from the figures for $2nk$, the error in n is also of the order of 2%. This error increases to about 5% in the region of 11 μ to 12 μ since errors in measurements of intensity become significant. The associated error in R_0 is $\frac{1}{4}$% or less, depending on the values of n and k; the figures for $R_0{}^2$ have therefore been given to three significant figures. The functions n^2-k^2 and $2nk$ are listed in the fourth and fifth columns of the table for convenience of reference since these are of greater value in interpretation than n and k. It may be noted that n^2-k^2 is large and negative; this is contrary to the assumption made by Hagen and Rubens (1904) in deriving an explicit expression for R_0.

Pressure, rate of evaporation and temperature at which the metal is condensed all influence the quality of mirrors of evaporated aluminium. On exposure to air an oxide film about 15 Å.U. thick forms very rapidly; subsequent growth of this layer lasts about a month when the oxide layer is about 40 Å.U. thick (Hass 1947, 1952). The optical constants therefore depend upon a number of factors; aluminium oxide (Al_2O_3) in the form of sapphire is not transparent at wavelengths greater than about 6 μ (Chasmer *et al.* 1951). For the present purposes of illustration a fuller examination of these effects is clearly out of place.

The polarization which is introduced by using such mirrors may be calculated using eqns. (5), (7) and (9) of the paper quoted above (Beattie and Conn 1954 a). If ϕ is small it follows from these equations when terms of the order of ϕ^4 are neglected that

$$1-\rho=2n\phi^2/(n^2+k^2) \text{ and } \Delta=2k\phi^2/(n^2+k^2).$$

Hence when ϕ is 4°, it follows that $1-\rho$ is of the order of 10^{-4} at 2 μ and Δ about 2 minutes of arc. These quantities decrease as the wavelength increases. It is clear therefore that aluminized mirrors may be employed with greater freedom than has perhaps been the custom when polarized infra-red radiation is employed.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge my gratitude to Dr. G. K. T. Conn for his encouragement and guidance. The benefit of many discussions with Dr. G. K. Eaton is also gratefully acknowledged.

This work has been supported by grants for apparatus from the Research Fund of the University of Sheffield and by the Royal Society.

REFERENCES

BEATTIE, J. R., and CONN, G. K. T., 1954, *Rev. Sci. Inst.*, **25,** 888; 1955, *Phil. Mag.*, **55** (in the press).

BOCK, R. C., 1945, *Phys. Rev.* **68,** 210.

BUECHE, F., 1948, *Journ. Opt. Soc. Am.*, **38,** 808.

CHASMAR, R. P., CRASTON, J. L., ISAACS, G., and YOUNG, A. S., 1951, *J. Sci. Inst.*, **28,** 206.

CONN, G. K. T., and EATON, G. K., 1954 a, *J. Opt. Soc. Amer.*, **44,** 546 ; 1954 b, *Ibid.*, **44**, 553.

CZERNY, M., 1924, *Zeits. f. Phys.*, **263,** 182.

ELLIOT, A., AMBROSE, E. J., and TEMPLE, R., 1948, *J. Opt. Soc. Amer.*, **38,** 212.

FÖRSTERLING, K., and FRÉEDERICKSZ, V., 1913, *Ann. d. Phys.*, **40,** 201.

HASS, G., 1947, *Anorg. Allgem. Chem.*, **254,** 96 ; 1952, *Vacuum*, **II,** 336 (pub. 1954).

HYDE, W. L., 1948, *J. Opt. Soc. Amer.*, **38,** 663 (Abs.).

INGERSOLL, L. R., 1910, *Astrophys. Jnl.*, **32,** 265.

JAMIN, J., 1847, *Ann. Chim. et Phys.*, **19,** 296.

LAÜE, M. VON, and MARTENS, F. F., 1907, *Verh. d. D. Phys. Ges.*, **5,** 522.

MCALISTER, E. D., MATHESON, G. L., and SWEENEY, W. J., 1941, *Rev. Sci. Inst.*, **12,** 314.

PFUND, A. H., 1933, *J. Opt. Soc. Amer.*, **23,** 275.

RUBENS, H., and HAGEN, E., 1904, *Phil. Mag.*, **7,** 162.

WILLIAMS, V. Z., 1948, *Rev. Sci. Inst.*, **19,** 135.

WILSON, A. H., 1936, *Theory of Metals* (C.U.P.) ; 1953, *Foreword to the Second Edition*.

XXVII. CORRESPONDENCE

The ' Boundary Effect ' in the Thomas–Fermi Model for Atoms

By R. A. Ballinger and N. H. March
Department of Physics, The University, Sheffield

[Received October, 1954]

The unmodified Thomas–Fermi (TF) method leads to a formula for the total binding energy of an atom which is not in good agreement with experiment. However, in an important paper, Scott (1952) has recently examined the reasons for the discrepancy, and has concluded that corrections must be applied for the so-called ' boundary effect ' and for exchange. One of us (March 1953) has examined the method employed by Scott for calculating the exchange energy, and has shown how a modified electron distribution, in which some account of exchange is introduced, can be set up. In this work, the ' boundary effect ' was not considered, and no modification of Scott's binding energy formula was called for, the only effect on the energy being the introduction of a term of order Z, whilst terms of order $Z^{4/3}$ are neglected in Scott's treatment.

The purpose of the present note is to make some remarks concerning the ' boundary effect ', and to suggest a modification to be made in Scott's formula. Scott has, effectively, estimated the correction to be applied to the TF formula by comparing the results of the TF approximation with the exact wave mechanical results for the case of particles moving in a Coulomb field. In the following we use a somewhat different method from that employed by Scott. If one considers Z particles, moving in a Coulomb potential Z/r (we use atomic units throughout), then the total energy when the lowest $\frac{1}{2}Z$ levels are doubly filled and we have n completed shells is

$$E=-Z^2n \qquad (1)$$

where n is related to Z by

$$\frac{n(n+1)(2n+1)}{6}=\frac{Z}{2}. \qquad (2)$$

We are now interested in an asymptotic form for E when Z is very large. It is easy to show from (2) that

$$n=(\tfrac{3}{2})^{1/3}Z^{1/3}-\tfrac{1}{2}+\tfrac{1}{18}(\tfrac{3}{2})^{2/3}Z^{-1/3}+O(Z^{-5/3}) \qquad (3)$$

and hence by substitution in (1) we have

$$E=-(\tfrac{3}{2})^{1/3}Z^{7/3}+\tfrac{1}{2}Z^2-\tfrac{1}{18}(\tfrac{3}{2})^{2/3}Z^{5/3}+O(Z^{1/3}). \qquad (4)$$

On the other hand, the TF approximation yields the result $-(\frac{3}{2})^{1/3}Z^{7/3}$ for all Z, and is therefore strictly correct only in the limit $Z \to \infty$ as one would expect from the nature of the method.

The term $\frac{1}{2}Z^2$ in (4) is just Scott's boundary correction, but we see that there is also now a term of order $Z^{5/3}$, and this is not negligible to the approximation considered by Scott. It seems clear to us from this derivation that to be consistent this term must be included in the binding energy formula. It is approximately one third of the exchange energy. In this way we obtain the modified formula

$$E = -0{\cdot}7687Z^{7/3} + \tfrac{1}{2}Z^2 - 0{\cdot}2936Z^{5/3} + O(Z^{4/3}). \quad . \quad . \quad . \quad (5)$$

We show in the table results for five atoms considered by Scott. For details of the way in which the ' observed ' values were obtained we refer the reader to Scott's discussion.

Binding Energies (Non-relativistic)

Z	6	9	13	18	23
' Observed '	37·8	99·7	242·4	528	945
Scott	36·7	97·6	236·9	518	933
Eqn. (5)	38·1	100·4	242·1	527	947

It will be seen that the results are in quite remarkable agreement with the observed values, a marked improvement being shown over the already satisfactory results obtained by Scott. In fact, the agreement is much better than one might expect from the nature of (5). However, the treatment we have given at least suggests a reason for this, for it will be seen from (4) that terms of order $Z^{4/3}$, Z, and $Z^{2/3}$ are absent. Whilst this will certainly not be the case with the formula (5), the results for the case of a Coulomb field make it seem very reasonable to suppose that these terms will be small, and hence that binding energies calculated from (5) will be more accurate than would appear at first sight.

In conclusion, we ought to point out that we have not of course thrown any light on the validity of the approximation implied in the use of a Coulomb field for the evaluation of the boundary correction. It does seem, however, that this is a less drastic approximation than one might imagine at first (for example, by noting that the first term of (4) is $-1{\cdot}14Z^{7/3}$, whereas that of (5) is $-0{\cdot}77Z^{7/3}$), for the bulk of the correction arises from the K-shell electrons, as was demonstrated by Scott. Concerning his derivation, the omission of the term in $Z^{5/3}$ which we have displayed here arises from the limit $n \to \infty$ which he takes, this procedure giving only the term of highest order in $Z^{1/3}$.

One of us (R. A. B.) wishes to acknowledge the award of an Ellison Fellowship by the University of Sheffield.

REFERENCES

MARCH, N. H., 1953, *Phil. Mag.*, **44**, 1193.
SCOTT, J. M. C., 1952, *Phil. Mag.*, **43**, 859.

XXVIII. REVIEWS OF BOOKS

The Physics of Particle Size Analysis. British Journal of Applied Physics Supplement No. 3. (London: Institute of Physics.) [Pp. 218.] Price, bound in cloth, 35*s*., post free.

In this volume there are three survey papers and thirty-three original contributions with the subsequent discussions which were presented at a conference on 'The Physics of Particle Size Analysis' held at the University of Nottingham in April 1954.

In recent years two new techniques have become widely used in particle size analysis: electron microscopy for the direct examination of fine dusts, and automatized counting and sizing with photoelectronic machines. Two sessions of the conference were devoted to the latter technique dealing firstly with the theory of automatic counting and sizing and secondly with the various photoelectronic instruments. Only one paper deals with the application of the electron microscope to particle size analysis.

At other sessions of the conference particle shape factors and visual counting and sizing of microscopic particles were discussed, whilst the physical properties of solid–liquid and solid–gas dispersions were fully treated in the first three sessions. It is rather surprising that not one paper deals with the properties and particle size analysis of solid–solid dispersions—a field of great interest to the petrologist and metallurgist.

The editors and organizing committee are to be congratulated on the publication of this volume so quickly after the conference.

J. N.

Journal of Nuclear Energy. (London: Pergamon Press Ltd.) [Pp. 300 approx. per volume. Quarterly at first.] Price £4 10*s*. 0*d*. per vol. for Libraries, £3 10*s*. 0*d*. for individuals.

The appearance of a new scientific journal is not always welcome for there are already so many journals to read. The *Journal of Nuclear Energy*, however, is an exception: the new and important branch of science and technology which it sets out to serve had hitherto no suitable outlet outside the United States. Nuclear energy projects are rapidly developing in many European and Commonwealth countries, and it is entirely fitting that there should be a European journal embracing all aspects of the work.

As the 'Atomic Age' gradually approaches the technologist begins to take precedence over the scientist. It is therefore to be expected that the first number of this Journal places its main emphasis upon the more technical aspects of nuclear engineering. Thus there are papers on the design of mechanical pumps for use with liquid metal coolants, on the control and operation of reactors, and on the economics of a large-scale power reactor programme. Nevertheless the more scientific aspects are by no means neglected. There are papers dealing with neutron cross section measurements and with the physical and physiological effects of radiation.

The international editorial board are to be congratulated upon the high standard which is set by their first number, the publishers upon the excellent quality of their print and plate. Engineers and scientists throughout the world will surely welcome the opportunity to collaborate within the pages of this Journal.

B. H. F.

[*The Editors do not hold themselves responsible for the views expressed by their correspondents.*]